TEMA

TRUSTED
EXTREMELY PRECISE
MAPPING AND PREDICTION
FOR EMERGENCY
MANAGEMENT

Deliverable D5.2: Report on algorithms for geovisual analytics and Digital Twin

TEMA
TRUSTED
EXTREMELY PRECISE
MAPPING AND PREDICTION
FOR EMERGENCY
MANAGEMENT

Project Information

Project acronym:	TEMA
Project full title:	Trusted Extremely Precise Mapping and Prediction for Emergency Management
Call identifier:	HORIZON-CL4-2022-DATA-01
Type of action:	HORIZON Research and Innovation Actions
Start date:	1 December 2022
End date:	30 November 2026
Grant agreement no:	101093003

Deliverable WP 5.2 - Geovisual Analytics			
Executive Summary:	**This deliverable describes the component called Geovisual Analytics. It reports the key feature and the role of this part inside the whole TEMA architecture. See the Section "Executive Summary".**		
WP:	**5.2**		
Author(s):	**See table below for a full list of authors**		
Editor:	**Cristian Federiconi**		
Leading Partner:	**Latitudo 40 SRL**		
Participating Partners:	**All**		
Version:	**1.0**	Status:	**Final**
Deliverable Type:	**R — Document, report**	Dissemination Level:	**PU - Public**
Official Submission Date:	**30/05/2025**	Actual Submission Date:	**10/06/2025**
Resubmission Date as requested by EC:	**09/04/2026**		

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

Disclaimer

The TEMA consortium consists of the following partners:

No.	Partner Organisation Name	Partner Organisation Short Name	Country
1	ARISTOTELIO PANEPISTIMIO THESSALONIKIS	AUTH	GR
2	DEUTSCHES ZENTRUM FUR LUFT – UND RAUMFAHRT EV	DLR	DE
3	ENGINEERING - INGEGNERIA INFORMATICA SPA	ENG	IT
4	ATOS IT SOLUTIONS AND SERVICES IBERIA SL	ATOS IT	ES
4.1	ATOS SPAIN SA	ATOS SP	ES
5	UNIVERSIDAD DE SEVILLA	USE	ES
6	TECNOSYLVA SL	TSYL	ES
7	NORTHDOCKS GMBH	ND	DE
8	PARIS-LODRON-UNIVERSITAT SALZBURG	PLUS	AT
9	THE LISBON COUNCIL FOR ECONOMIC COMPETITIVENESS ASBL	LC	BE
10	LATITUDO 40 SRL	LAT40	IT

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

11	NELEN & SCHUURMANS TECHNOLOGY BV	NS	NL
12	FRAUNHOFER GESELLSCHAFT ZUR FORDERUNG DER ANGEWANDTEN FORSCHUNG EV	FHHI	DE
13	UNIVERSITA DEGLI STUDI DI MESSINA	UNIME	IT
14	KAJAANIN AMMATTIKORKEAKOULU OY	KAMK	FI
15	KAJAANIN KAUPUNKI	KAJ	FI
16	KENTRO MELETON ASFALEIAS	KEMEA	GR
17	DIMOS MANTOUDIOU - LIMNIS - AGIAS ANNAS	D.MALLIAN	GR
18	REGIONE AUTONOMA DELLA SARDEGNA*RAS	RAS	IT
19	BAYERISCHES ROTES KREUZ	BRK	DE
20	KAINUUN HYVINVOINTIALUE	KAHY	FI
21	INTERDISCIPLINARY TRANSFORMATION UNIVERSITY AUSTRIA	IT:U	AT

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

Document Revision History

Version	Description	Contributions
0.1	ToC	Cristian Federiconi (LAT40) Giovanni Giacco (LAT40) Dmitriy Shutin (DLR)
0.2	First Draft	Cristian Federiconi (LAT40) Giovanni Giacco (LAT40) Dmitriy Shutin (DLR)
1.0	Final Version	Cristian Federiconi (LAT40) Giovanni Giacco (LAT40) Dmitriy Shutin (DLR)

Authors/contributors

Author	Partner
Cristian Federiconi	LAT40
Giovanni Giacco	LAT40
Vasileios Mygdalis	AUTH
Ioannis Pitas	AUTH
Michael Siavrakas	AUTH
Eugenios Vlachos	AUTH
Joachim Perschbacher	ND
Viktor Bezzmertnyi	KAMK

Reviewers

Name	Organisation
Dmitriy Shutin	DLR-KN

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

Table of Contents

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

TEMA
TRUSTED
EXTREMELY PRECISE
MAPPING AND PREDICTION
FOR EMERGENCY
MANAGEMENT

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

List of Figures

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

List of Tables

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

List of Terms and Abbreviations

Abbreviation	Description
TEMA	Trusted extremely precise mapping and prediction for emergency management
OSM	OpenStreetMap
COG	Cloud Optimized GeoTIFF
ETL	Extract, Transform, and Load
API	Application Programming Interface
GHCR	GitHub Container Registry
CPU	Central Processing Unit
RAM	Random Access Memory
SDK	Software Development Kit
CRS	Coordinate Reference System
NGSI-LD	Next Generation Service Interface - Linked Data
WMS	Web Map Service
KPI	Key Performance Indicator
XR	Extended Reality

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

Executive Summary

Deliverable D5.2 presents a comprehensive report on the Geovisual Analytics component (Task T5.2), a pivotal element within Work Package 5 "Simulation and Visualization" of the TEMA project. The Geovisual Analytics module is engineered to facilitate the rapid retrieval, processing, enrichment, and preparation of large-scale, heterogeneous geospatial data, transforming it into actionable intelligence for Natural Disaster Management (NDM). It plays a crucial intermediary role, bridging foundational 3D environmental models from the Precise Digital Twin (T5.1) with the advanced interactive visualization interfaces of the SmartDesk and XR Viewer (T5.3).

Key functionalities of the Geovisual Analytics component include advanced **Geospatial Information Retrieval** for extracting relevant data, **AI-driven Semantic Annotation** for enhancing geospatial metadata and searchability, and a suite of **Geospatial Computation & Analytics** capabilities. These core operations encompass OpenStreetMap (OSM) searches for location-based queries, efficient GeoTIFF to Cloud Optimized GeoTIFF (COG) conversion, critical distance-to-nearest point calculations (e.g., "distance from fire"), and zonal statistics for extracting summary information. The component is validated through real-world NDM scenarios, specifically fire and flood risk geovisual analytics, with results made accessible through **Dashboard Integration** with the TEMA SmartDesk.

Architecturally, the Geovisual Analytics module features an event-driven ETL (Extract, Transform, Load) pipeline. It employs a modular design with reusable "Analytics Building Blocks" and leverages **Dask** for parallel processing and workload optimization, ensuring scalability and performance when handling "extreme data" volumes. Seamless integration with external systems and other TEMA components is achieved via a **REST API (FastAPI)** and, most importantly, through the **FIWARE Context Broker**. Adherence to standardized FIWARE data models is a cornerstone, ensuring semantic interoperability across the TEMA ecosystem. This allows the component to consume diverse inputs effectively and publish enriched, context-aware geospatial products for downstream visualization and decision-support.

Continuous improvements are applied based on iterative refinement and feedback, enhancing data accuracy, processing speed, and usability. Ultimately, the Geovisual Analytics component serves

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

as a crucial enabler within TEMA, transforming raw and modeled geospatial data into timely, context-rich, and semantically interoperable insights, thereby significantly enhancing situational awareness and decision-making capabilities for emergency responders.

1 . Introduction

1.1 Purpose and scope of the document

This document, **Deliverable D5.2 "Report on algorithms for geovisual analytics and Digital Twin,"** details the **Geovisual Analytics component** developed within **Task T5.2** of TEMA's Work Package 5. It outlines the component's objectives, architecture, core algorithms, data processing workflows, and its crucial role in the TEMA ecosystem.

The primary **scope** is the Geovisual Analytics module, covering its function as an intermediary engine that processes and enriches geospatial data, including outputs from the Precise Digital Twin (T5.1). A key focus is its integration with the TEMA platform via FIWARE for semantic interoperability, preparing data for visualization by the SmartDesk and XR Viewer (T5.3). This report highlights how T5.2 handles large-scale data using `Dask` for parallelization, and discusses its deployment and testing. While contextualizing its interactions with T5.1 and T5.3, detailed algorithmic specifics of those tasks are reserved for Deliverable D5.3. This document underscores T5.2's contribution to the TEMA proof-of-concept NDM system by enabling effective and insightful visualization.

1.2 Structure of the document

This document is organized to provide a clear understanding of the Geovisual Analytics component within TEMA WP5. **Section 2** summarizes the work carried out, outlining the specific objectives of Task T5.2 and detailing the progress made in achieving them, particularly in relation to the broader WP5 goals. **Section 3** offers a general introduction to Geovisual Analytics, discussing its definition, challenges in creating effective systems, and the current state-of-the-art. **Section 4** delves into Geovisual Analytics specifically within the TEMA project. It presents an overview of the module, its functional objectives, how it addresses key challenges, its integration with other TEMA visualization components (Digital Twin, SmartDesk, XR Viewer), and a summary of the work

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

undertaken. **Section 5** details the core components of the Geovisual Analytics module, including its architecture, data processing workflows, interfaces with other TEMA components (with a focus on FIWARE integration), key building blocks like the "distance from fire" analytic, REST API design, geospatial image retrieval, and considerations for deployment and scalability. **Section 6** concludes the report, summarizing key achievements and outlining the future roadmap for the Geovisual Analytics component.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

2 Summary of the work carried out

2.1 Objectives

TEMA addresses the challenges in Natural Disaster Management (NDM) for events like regional floods, flash floods, and wildfires by leveraging heterogeneous data sources—including edge devices, sensors, satellite imagery, geospatial data, meteorological information, and geosocial media—to provide enhanced simulation and visualization capabilities. These data sources are inherently complex, voluminous, and dynamic. The main objective of **TEMA WP5 "Simulation and Visualization"** is to develop novel, intuitive, and interactive interfaces that fuse a rapidly constructed geovisual map with precise Digital Twin representations. This fusion will empower NDM human operators to immersively review current situations, assess near-future predictions and recommendations generated by other TEMA work packages (WP3, WP4), and evaluate contingent response alternatives, thereby facilitating the selection of optimal courses of action.

The specific TEMA objectives centrally addressed by **Task T5.2 "Geovisual analytics"** (and supported by the broader WP5 context) are derived from the critical need to efficiently process, analyze, and prepare vast amounts of geospatial data for these advanced visualization environments. These objectives, along with their accompanying Key Performance Indicators (KPIs) and Target Values (TVs) as defined in Section 1.1.1 of Part B of TEMA's Description of the Action (DoA) and updated in project reviews (M24 KPI Status), include:

- **OB3) Improve responsiveness and interactivity of visualization mechanisms for evolving phenomena (WP5):** Task T5.2 directly contributes to this by ensuring the rapid processing and delivery of geovisual data to the visualization tools developed in T5.3. The target KPI includes reducing "Data-to-visualization time up to 30%, compared to SoA geovisual analytics systems" and ensuring "The response time of flood models will be near-realtime."
- **OB4) Improve accuracy of visualization mechanisms for evolving phenomena (WP5):** By providing accurately processed, semantically enriched, and high-resolution geospatial maps, Task T5.2 enhances the fidelity and detail of the visualizations. The target KPIs,

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

"Digital Twin accuracy (TV: Reduction of RMSE up to 30% compared to SoA for identical runtime requirements)" and "Merged spatial 3D map resolution (TV: Increase up to 30% over SoA for identical runtime requirements)," have both been reached.

- **OC4) Prototype a proof-of-concept TEMA system for NDM in forest fires, flash floods, and regional floods (WP2-6):** The Geovisual Analytics module (T5.2), in conjunction with T5.1 and T5.3, forms a significant part of this proof-of-concept system. By processing, integrating, and enabling the visualization of relevant geospatial data for these specific disaster types, T5.2 is instrumental in demonstrating the end-to-end functionality of the TEMA NDM system. The successful integration and visualization of fire and flood-related analytics in the SmartDesk are direct contributions to this objective.
- Task T5.2 also indirectly supports **OC2) Increase situational awareness in NDM (WP2-6)** by preparing and structuring the foundational geospatial data that is visualized by the end-user interfaces.

2.2 Summary of the work carried out with respect to the objectives

The activities within WP5 are geared towards creating a seamless and efficient pipeline, transforming raw and processed geospatial data into interactive, insightful visualizations tailored for NDM operators. While this deliverable (D5.2) concentrates on the pivotal intermediary role and developments within **Task T5.2: Geovisual analytics**, we also briefly present here the main advancements in Digital Twin creation (Task T5.1) and the specific features of the XR Viewer and Smartdesk interfaces (Task T5.3) until this period. The detailed results will be comprehensively described in Deliverable D5.3 ("Interactive Digital Twin for critical emergencies"),

- **T5.1: Precise Digital Twin (ND, ENG, ATOS; Led by Northdocks):** This task focuses on developing novel photogrammetry solutions for generating precise Digital Twins of afflicted areas in real-time. Progress includes significant work with historical sample data, exploration of algorithms, definition of needed input data (particularly drone images), and the creation of Deliverable D5.1 ("Report on algorithms for

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

precise Digital Twin construction"). Integration is currently in progress, following the established integration plan. The outputs of T5.1, such as the "Model3D" from the SV03 Gateway, provide the fundamental 3D environmental context.

- **T5.2: Geovisual analytics (LAT40, AUTH, ENG, TSYL, NS, DLR; Led by Latitudo40):** This task, develops a novel cluster-based system for the rapid retrieval, processing, and semantic annotation of geospatial data. Progress will be detailed in this deliverable, with key achievements including the following:
    - **FIWARE Integration and Event-Driven Processing:** Successful implementation of understanding the FIWARE entity model managed in the Context-Broker, enabling subscriptions to entities of interest (e.g., "Map4Fire"). This triggers the Geovisual Analytics processes, such as calculating the "distance from fire" for buildings in the Fire Business Model.
    - **Data Output and Storage:** Processed output data is stored efficiently in Minio and relevant context is updated in the Context-Broker, ensuring data persistence and accessibility for other TEMA components.
    - **Parallelization and Workload Balancing (Dask):** To handle the "extreme data" nature described in the TEMA proposal, algorithms leverage Dask for parallel processing of geospatial data. This involves breaking down large datasets (rasters, vectors) into manageable chunks for distributed computation, optimizing for speed and resource utilization, especially within the Kubernetes cluster environment.
    - **Semantic Annotation and Enrichment:** Building upon analytical outputs from WP3 (e.g., TFA11 Geosocial media analysis) and WP4 (e.g., PDM-tech-05 Information fusion), T5.2 is responsible for integrating these insights into a cohesive geospatial map. This involves semantically annotating raw geospatial data and the Digital Twin outputs from T5.1.
    - **Optimized Data Formats and FIWARE Interoperability:** A core function is the transformation of diverse inputs into standardized FIWARE data models, ensuring semantic interoperability. This processed and enriched data (e.g., "DistanceFromFire") is then published for consumption by visualization tools.
    - **Algorithm Refinement:** Continuous refinement of algorithms, such as filtering for vulnerable and important buildings (bridges, hospitals), is conducted based on feedback from partners, including KAMK for data visualization aspects.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

- **T5.3: Augmented Reality and rapid visualization (ND, DLR, LC, NS, TSYL, FHHI, KAMK; Smartdesk led by KAMK, XR Viewer by Northdocks):** This task develops the novel, real-time XR-based interactive visualization prototype. The Smartdesk (KAMK) has shown considerable progress, incorporating improvements from hackathons and end-user feedback. It now capably displays alerts, emergency maps, UAV trajectories, and 3D map entities (derived from T5.1 and processed/served via T5.2), and features dedicated "Geo Analytics" cards that directly visualize T5.2 outputs. A service to connect the Smartdesk with the ORION Digital Enabler (Context Broker) for real-time data and support for signed MinIO URLs has also been implemented. The XR Viewer (Northdocks) has completed the implementation of file loaders for various TEMA partner data formats (Cesium Tiles, KML, GeoJSON) and is in the planning and requirements development phase for its first version, with UI development proceeding in collaboration with end-users.

The integrated workflow within WP5 clearly shows T5.1 providing the foundational 3D environmental models. Task T5.2 then acts as the crucial analytical and preparatory engine, processing and enriching these models along with other dynamic geospatial inputs, ensuring semantic consistency via FIWARE standards, and optimizing for rapid delivery. Finally, Task T5.3 components (Smartdesk and XR Viewer) consume these readily digestible, context-rich geospatial products from T5.2 to provide end-users with powerful and intuitive visualization tools for effective NDM.

3 . Geovisual Analytics

3.1 Introduction to Geovisual Analytics

Geovisual analytics represents an interdisciplinary field dedicated to analytical reasoning facilitated by interactive visual interfaces, with a specific focus on geospatial data and phenomena. It emerges from the confluence of Geographic Information Science (GIS), information visualization, cartography, data mining, and human-computer interaction. The fundamental premise of geovisual analytics is to empower human analysts to gain deeper insights from complex, voluminous, and

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

often dynamic geospatial datasets that might otherwise remain obscured through traditional analytical methods or static visualizations alone.

A widely accepted formal definition posits geovisual analytics as "a research field that combines the strengths of human and machine capabilities for effective and efficient reasoning with geospatial information by means of interactive visual interfaces" [1][2]. It emphasizes the synergy between automated data analysis techniques and the perceptual and cognitive capabilities of human analysts to extract insights from complex spatial and spatio-temporal datasets.

At its core, geovisual analytics entails the seamless integration of automated computational data processing techniques with the unparalleled perceptual and cognitive capabilities of human users. This synergy is achieved through the transformation of raw geospatial data into meaningful visual representations—such as dynamic maps, linked statistical graphics, spatio-temporal animations, and three-dimensional scenes. These representations are not merely for presentation but are designed to be actively explored. Users can interactively query data, adjust visualization parameters, filter information, and navigate through multiple scales and perspectives, thereby fostering an exploratory data analysis environment. This interactive exploration is crucial for identifying spatial patterns, detecting anomalies, understanding relationships, formulating hypotheses, and ultimately, communicating findings effectively.

In the context of complex problem domains such as emergency management, urban planning, environmental monitoring, and public health, geovisual analytics provides powerful mechanisms. It supports the synthesis of diverse and heterogeneous data streams (e.g., sensor networks, remote sensing imagery, social media feeds, simulation outputs, demographic statistics) into a coherent and comprehensible operational picture. By enabling stakeholders to visually explore and analyze these integrated datasets, geovisual analytics enhances situational awareness, facilitates pattern detection, supports collaborative problem-solving, and aids in informed decision-making under conditions of uncertainty and time pressure. The importance of this field lies in its ability to bridge the gap between the increasing availability of geospatial data and the human capacity to derive actionable intelligence from it, particularly when addressing multifaceted real-world challenges.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

3.2 Challenges in Creating Effective Geovisual Analytics Systems

Creating effective geovisual analytics systems—those that combine computer power with human thinking to make sense of spatial data—faces several complex and interconnected challenges. These challenges span the entire process, from handling complex data and designing efficient systems, to ensuring smooth user interaction and evaluating how well the system works. Overcoming these issues is essential to improve the usefulness and impact of geovisual analytics in many real-world applications.

At the forefront of these challenges are those stemming from the nature of contemporary geospatial data. Modern datasets are frequently characterized by what is often termed "Big GeoData" attributes, presenting substantial hurdles in terms of volume, velocity, and variety [3, 4]. Geovisual analytics systems must therefore be engineered to manage and process immense quantities of information, often arriving in dynamic, real-time streams and originating from a diverse array of heterogeneous sources, each with its own unique formats, structures, and semantic underpinnings. Beyond these foundational data management issues, the reliability of analytical outcomes is intrinsically tied to data quality, the handling of inherent uncertainty, and the transparent tracking of data provenance. Geospatial information is often imperfect, exhibiting missing values, positional inaccuracies, or attribute errors. Consequently, a significant challenge lies in developing robust methods to not only address these imperfections but also to effectively represent and communicate the pervasive uncertainty within both the raw data and the derived analytical products to the end-user, thereby guarding against misinterpretation and supporting more nuanced decision-making [5, 6]. Furthermore, the task of integrating data from disparate sources to create a cohesive analytical view often requires overcoming significant semantic heterogeneity and ensuring interoperability between different data models and technological systems, a pursuit that increasingly involves sophisticated semantic web technologies and knowledge graph approaches [7, 8]. Compounding these data-centric issues are the growing ethical and technical imperatives surrounding privacy and security, particularly as geovisual analytics leverages detailed, often person-specific, location-based information, necessitating the incorporation of robust privacy-enhancing technologies without unduly constraining analytical exploration [9, 10].

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

From a computational and system design perspective, the demand for scalability and interactive performance remains a persistent and critical consideration. The core value proposition of geovisual analytics rests on its ability to support fluid, interactive exploration by the user, which translates directly into a stringent requirement for near real-time system responsiveness to queries and manipulations [11]. Achieving this level of interactivity when dealing with voluminous datasets and computationally intensive spatial analyses—such as sophisticated geostatistical modeling, network traversal algorithms, or complex simulations—frequently necessitates recourse to advanced algorithmic optimization, efficient data indexing strategies, and often the adoption of high-performance computing paradigms, including distributed and cloud-based architectures [12, 13]. A pivotal design consideration in this context is the tight coupling of computational analysis with interactive visualization, a symbiotic relationship where analytical algorithms dynamically inform visual representations, and conversely, user interactions with visualizations parameterize or trigger further computational analyses [14, 15]. Architecting systems that facilitate this bidirectional flow of information and control in an efficient and flexible manner continues to be an active area of research and development. Moreover, the long-term viability and adaptability of geovisual analytics systems are greatly enhanced by a modular design approach that promotes the development of reusable components for data handling, analysis, and visualization. However, this desirable modularity introduces its own set of complexities, particularly in defining robust Application Programming Interfaces (APIs), managing inter-component dependencies, and ensuring seamless integration across diverse software modules [16].

3.3 State-of-the-art in Geovisual Analytics

The current geovisual analytics landscape includes several mature categories of solutions: desktop GIS environments, web-based geospatial publication platforms, large-scale cloud analytics services, real-time event analytics platforms, and 3D geospatial digital twin frameworks.

Several open-source and commercial platforms have emerged to support these capabilities. QGIS, for instance, combined with a robust spatial relational database management system (RDBMS) like PostgreSQL/PostGIS, provides a highly capable enterprise stack for database-backed spatial querying, editing, and visualization. With appropriate server hardware and performance tuning, a PostGIS-backed architecture can manage millions of concurrent spatial records and complex topological queries in real-time. Similarly, GeoNode supports web-based management and

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

publication of geospatial content, while Google Earth Engine provides planetary-scale analysis over large remote sensing archives. In the commercial sphere, ArcGIS GeoEvent Server and ArcGIS Velocity provide real-time ingestion, transformation, geofencing, and alerting for event streams, and Cesium-based platforms deliver high-quality 3D geospatial visualization for tiled 3D content.

It is therefore important to state explicitly that TEMA is not claiming novelty in basic geospatial visualization, spatial SQL, database-backed map rendering, or generic 3D scene rendering. In particular, QGIS and PostGIS already cover a substantial portion of conventional geospatial analysis workflows and can easily support server-backed, large-scale processing when properly deployed.

The added value of TEMA must instead be positioned in the way analytics are orchestrated, triggered, semantically exposed, and consumed across a highly specialized platform. The innovation does not lie in replacing general-purpose GIS tools, but in creating an event-driven, AI-native geovisual analytics capability specifically tailored to emergency-management workflows. A requirement-driven analysis shows why existing tools, while powerful, do not completely satisfy the specific objectives of TEMA out of the box. Developing a custom geovisual analytics pipeline for TEMA fulfills unique architectural imperatives that mature GIS platforms cannot natively support without massive, complex custom engineering:

- **Native AI and Tensor Integration:** Traditional spatial databases (like PostGIS) are not built for deep learning pipelines. TEMA relies heavily on Deep Neural Networks (DNNs) for extreme data analytics, such as the Concept-Based Content-Based Image Retrieval (CCBIR) pipeline, Explainable AI (XAI), and multimodal data fusion. TEMA's custom architecture—leveraging Python-native distributed frameworks like Dask—allows for the processing of tensors, raster arrays, and neural network inferences within the exact same computational pipeline. Attempting to force these operations through a traditional SQL database creates severe I/O bottlenecks and data serialization overheads.
- **Event-Driven Semantic Interoperability (FIWARE/NGSI-LD):** Traditional GIS relies heavily on static database polling or user-driven visual interaction models via standard OGC web services (WMS/WFS). TEMA requires automatic reaction to context updates via a publish/subscribe approach, orchestrated by the FIWARE Context Broker using the NGSI-LD standard. This allows the geovisual analytics module to instantly trigger downstream analytical chains the moment a drone uploads a new image or a fire simulation updates, ensuring semantic interoperability across heterogeneous IoT sensors and AI modules.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

- **Mission-Oriented Heuristics:** A distinctive element of the TEMA approach is the heuristic layer. Instead of exposing all derived spatial layers with the same weight, TEMA utilizes a mission-oriented set of rules and ranking criteria that translates raw geospatial outputs into operationally meaningful priorities. This heuristic combines hazard proximity, asset criticality, temporal evolution, and operator-defined thresholds to foreground what is most relevant for action. In emergency management, visual completeness alone does not guarantee actionable situational awareness.
- **Inflexibility across the Edge-to-Cloud Continuum:** NDM operations frequently occur in disaster zones where network connectivity is degraded. Monolithic database-driven GIS architectures struggle with dynamic, federated distribution. TEMA's containerized, microservice-based Analytics Building Blocks are explicitly designed to be deployed flexibly—either in a centralized cloud or pushed to the "edge" (e.g., executing directly on a drone ground station or a SmartDesk within a local command vehicle).
- **Friction in Direct Feed to XR and Precise Digital Twins:** The output of TEMA's geovisual analytics is designed to feed immediately into immersive 3D Augmented Reality (XR) environments and highly precise Digital Twins (T5.1 & T5.3). The TEMA pipeline is optimized to transform data into specific, lightweight, and semantically enriched formats tailored explicitly for AR rendering engines.

The comparison in *Table* Table 1 should therefore be read as a positioning exercise rather than as a binary ranking of tools. Existing platforms remain highly valuable and, in some cases, are complementary to TEMA. However, none of them natively matches the exact combination of requirements pursued in TEMA.

Table 1– Comparative positioning of the TEMA Geovisual Analytics module against representative state-of-the-art solutions

Platform / Stack	***Main Strength***	***Relevance for TEMA***	***Why it is not sufficient on its own for TEMA***
QGIS + PostgreSQL/ PostGIS	Mature desktop GIS visualization, database connectivity, spatial SQL, indexed spatial queries, analyst-driven workflows.	Strong benchmark for conventional GIS analysis and visualization.	Very strong for expert-driven GIS work, but not conceived as a FIWARE-native event-processing node that subscribes to context updates, integrates native DNN

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

			inferencing, and distributes results to multiple consumers in a closed loop.
GeoNode	Web-based geospatial data management, publication, sharing, SDI-style dissemination.	Relevant for data publication and cataloguing.	Useful for publishing and access, but not primarily designed for mission-oriented, event-triggered analytics orchestration in emergency workflows.
Google Earth Engine	Planetary-scale geospatial processing over large Earth Observation (EO) archives.	Relevant as a reference for scalable remote sensing analysis.	Very strong for large-scale EO computation, but not designed as the operational integration layer for TEMA's FIWARE-based event flows, local Edge deployment, or heuristic filtering.
ArcGIS GeoEvent / ArcGIS Velocity	Real-time feeds, transformation, geofencing, enrichment, alerts.	Relevant commercial benchmark for event analytics.	Covers real-time analytics well, but remains a proprietary commercial ecosystem and does not represent TEMA's open, FIWARE-centred, AI-integrated building-block architecture.
Cesium-based digital twin platforms	High-quality 3D geospatial rendering, 3D Tiles streaming, digital twin visual context.	Relevant benchmark for 3D operational visualization.	Excellent for 3D rendering and scene interaction, but not in itself the analytic orchestration layer that computes, ranks, and semantically republishes mission outputs across services.
TEMA Geovisual Analytics	Event-driven FIWARE integration, native AI/Tensor processing, reusable building blocks, edge-to-cloud	Purpose-built for TEMA business missions.	Not intended to replace all existing GIS products; its role is to integrate, temporalize, enrich, prioritize, and distribute AI-driven geospatial analytics inside the TEMA architecture.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

	deployment, mission-oriented heuristics.

Note: The capability summaries in Table Table 1 *are grounded in the official documentation of the referenced platforms and should be interpreted together with the specific requirements of the TEMA architecture rather than as a generic market ranking.*

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

4. Geovisual Analytics in TEMA

This chapter introduces geovisual analytics, a field that combines spatial data analysis with interactive visualization to support decision-making. It explores its key components, challenges, and applications, particularly in the context of emergency management.

4.1 Overview of the Geovisual Analytics Module in TEMA

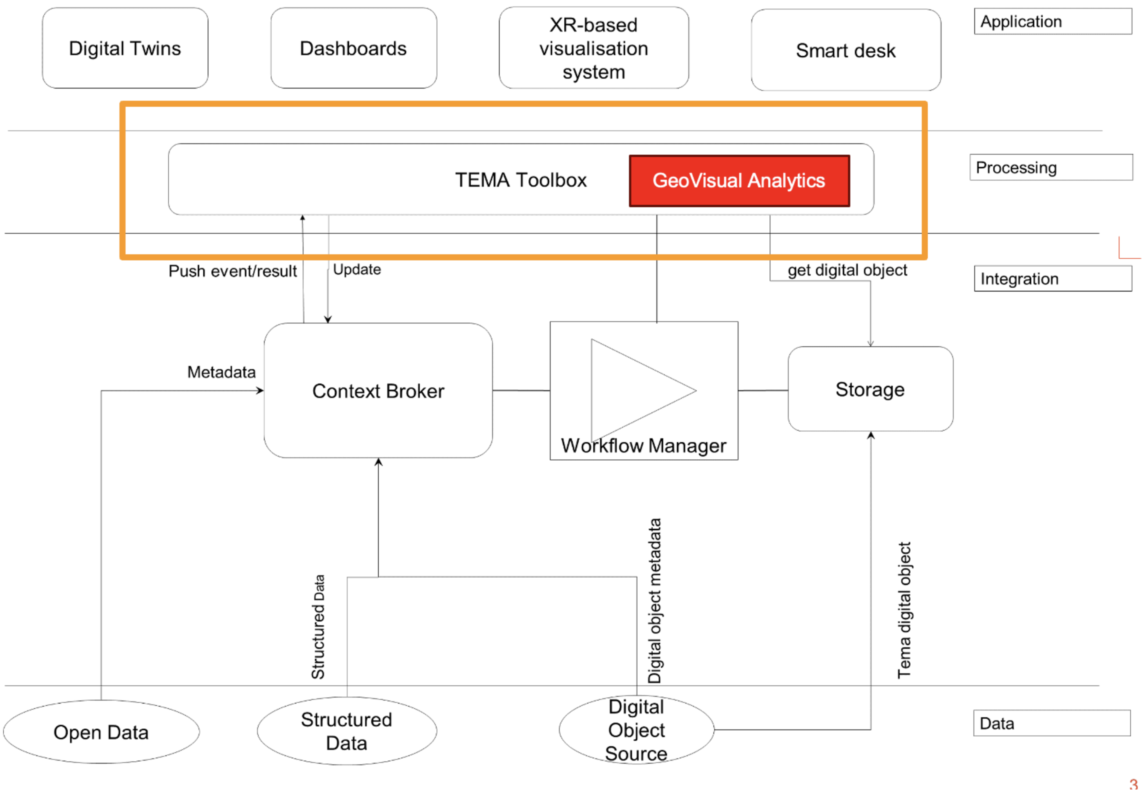


Figure 1. Geovisual Analytics Toolbox in the TEMA Platform

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

The Geovisual Analytics module is a critical, event-driven processing component within the TEMA platform, specifically situated as a core element of the TEMA Toolbox, as illustrated in the TEMA platform's architectural diagram (Figure 1). Its primary function is to efficiently ingest, extract, transform, load (ETL), analyze, and enrich diverse, large-scale geospatial data from various sources—including Open Data, Structured Data, and Digital Object Sources managed via the FIWARE-based Context Broker (digital enabler) and Workflow Manager. By integrating additional contextual information, such as from OpenStreetMap, the Geovisual Analytics module transforms raw datasets into actionable intelligence and semantically enriched information products.

These products are fundamental for supporting the TEMA application layer, most notably the Digital Twins, Dashboards, the XR-based visualisation system, and the Smart desk. The Geovisual Analytics module employs advanced computational techniques, leveraging parallel and distributed processing libraries like `rasterio`, `rioxarray`, and `Dask`, to handle complex raster and vector data operations at scale. Key functionalities employed within TEMA include spatial queries on OpenStreetMap data to identify relevant points of interest, conversion of raster datasets into cloud-optimized formats for efficient access, computation of zonal statistics for defined areas, and distance calculations between critical features and event locations.

The Geovisual Analytics module's architecture is modular, based on reusable analytics building blocks, allowing for straightforward expansion and integration of new functionalities without disrupting existing workflows. Designed for deployment in containerized cluster environments such as Kubernetes, it interacts with the Context Broker through event notifications and subscriptions to orchestrate data flow and processing in (near) real-time. Emphasis is placed on optimized code execution for near real-time performance and comprehensive logging, favoring lightweight and maintainable solutions.

By providing these processed and analyzed geospatial insights ("digital objects") in formats optimized for visualization and spatial analysis, the Geovisual Analytics module directly enables the TEMA platform to achieve its objectives of precise mapping, phenomenon prediction, and effective decision support for Natural Disaster Management (NDM), including fire and flood risk assessment. Its outputs are crucial for updating the situational picture, feeding into predictive models, and providing the foundational geospatial context for interactive visualizations and the Digital Twin, thereby enhancing situational awareness and supporting timely, accurate, and

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

context-rich emergency response. Overall, geovisual analytics serves as a critical enabler within the TEMA project's geospatial data ecosystem, transforming heterogeneous spatial data into actionable intelligence through scalable, modular, and efficient analytical workflows.

4.2 Functional Objectives of the Geovisual Analytics in TEMA

The Geovisual Analytics component has been conceived with a set of functional objectives that collectively ensure the delivery of timely, accurate, and context-rich geospatial intelligence within the TEMA platform.

A primary objective is to enable **seamless data ingestion, harmonization, and semantic interoperability**. This involves the extraction, transformation, and loading (ETL) of heterogeneous spatial data, ensuring diverse raw inputs become immediately usable for situational awareness. Data arriving from satellites, drones, open data portals, field sensors, and partner technologies must be harmonised, spatially referenced, and crucially, enriched with semantic metadata. This objective directly addresses semantic interoperability by standardizing data representation across the TEMA platform. This is fundamentally achieved through the adoption and implementation of FIWARE data models, which provide a common, standardized structure for geospatial information. By adhering to these models, the Geovisual Analytics module ensures that data pushed to and consumed via the FIWARE Context Broker is consistently understood by all TEMA components (including the Digital Twin, SmartDesk, and XR visualization), facilitating seamless integration and meaningful data exchange. This ETL and semantic enrichment capability is implemented through a chain of reusable operations rigorously profiled for execution speed and memory efficiency, as emergency-management scenarios cannot tolerate lengthy pre-processing stages.

Another key functional objective concerns **advanced analytical depth for mission-specific insights**. The component must support complex spatial queries such as point-in-polygon tests, proximity computations, zonal statistics over arbitrary raster masks, and multi-temporal change detection. These operations are critical for the fire, flood, and other emergency management business missions defined in TEMA. To reach the analytical depth demanded by these use cases, the component exposes composable building blocks implemented on top of libraries like rasterio, rioxarray, and dask, thereby allowing practitioners to chain together fine-grained functions efficiently without the overhead of heavyweight workflow orchestrators.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

The provision of **rapid and interactive visual feedback** is also a critical functional objective. Insights produced by Geovisual Analytics must travel quickly to SmartDesk dashboards and external Digital Twin viewers, as decision-makers often refine their information needs on the fly. Therefore, the component provides a low-latency publish-subscribe interface, based on the FIWARE Context Broker, which pushes semantically interoperable results directly to consumer services. Response times are continuously benchmarked: every processing cycle is measured from notification receipt to object-storage write, then from storage write to context-broker metadata update, with goals set to keep the end-to-end latency below the thirty-second mark under normal network conditions.

Furthermore, an important objective lies in **scalability** for varied disaster scopes. Natural disasters can span from a single municipality to an entire region. The component must thus be able to expand its processing capacity from a single-node test deployment to a multi-node Kubernetes cluster without requiring code changes. Parallel execution is achieved by partitioning raster tiles and vector feature batches across threads and workers managed by libraries such as dask. Memory consumption is kept proportional to tile size, and an adaptive chunking policy ensures that partial results are flushed to object storage as soon as they are produced, preventing memory saturation when extremely large rasters are involved.

Reliability and provenance tracking constitutes another essential objective. Each processing run is wrapped in an atomic transaction that writes a provenance record—including the hash of all input artefacts, the version of each software dependency, and the list of parameter values—into the FIWARE Context Broker. Should a pod crash, an idempotent replay mechanism evaluates these records and restarts the unfinished step, guaranteeing that results are never silently corrupted. Extensive logging at debug, info, and error levels is shipped to a centralised observability stack, enabling operators to trace anomalies down to individual spatial tiles or vector batches.

Finally, **extensibility and maintainability** are core objectives. The design favors modularity over monolithic growth. Every new analytical capability must be introduced as an independent building block with a well-defined interface signature. Versioning rules enforce backward compatibility at the API boundary, permitting external systems to adopt new analytics without requiring code freezes. Documentation is generated automatically from inline type hints and docstrings, ensuring that each function clearly advertises accepted geometries, coordinate-reference systems, and expected output formats.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

These functional objectives, with a strong emphasis on semantic interoperability through FIWARE data models, together provide the foundation for a Geovisual Analytics service throughout the TEMA project lifecycle.

4.3 Requirements-Driven Gap Analysis and TEMA Added Value

The objective of this section is not to claim that existing geospatial tools are unable to visualize disaster-related outputs. On the contrary, many mature solutions—such as QGIS combined with PostgreSQL/PostGIS, ArcGIS GeoEvent Server, or Google Earth Engine—already provide high-quality visualization, database connectivity, spatial querying, and even real-time event analytics. QGIS and PostGIS, in particular, cover a substantial portion of conventional geospatial analysis workflows and can easily support server-backed, large-scale processing when properly deployed.

The relevant question for TEMA is more specific: whether those solutions, taken individually or in standard combinations, satisfy the *full* set of operational and architectural requirements of the project. It is important to state explicitly that TEMA is not claiming novelty in basic geospatial visualization, spatial SQL, or generic 3D scene rendering. The innovation does not lie in replacing general-purpose GIS tools with a better generic viewer, but in creating an event-driven, AI-integrated geovisual analytics capability specifically tailored to emergency-management workflows.

A requirement-driven analysis shows why existing tools, while powerful, do not completely satisfy the specific objectives of TEMA out of the box, necessitating the development of a custom, composable pipeline:

- **Event-Driven Operation and Semantic Interoperability:** In TEMA, geovisual analytics is not only expected to display results after a user request; it must react automatically to incoming context changes, trigger the relevant processing chain, and publish derived outputs back into the platform. Desktop GIS tools can visualize updated data effectively, but they are not designed to be the core *event-driven integration node* of an autonomous platform. TEMA's event-driven backbone is enabled by the **FIWARE Context Broker** and the NGSI-LD standard. This publish/subscribe logic allows analytics to be activated

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

instantly by context updates (e.g., a drone uploading an image) and redistributes the results as machine-readable, semantically interoperable entities to multiple consumers (SmartDesk, XR interfaces).

- **Native AI and Tensor Integration:** Traditional spatial databases (like PostGIS) excel at vector mathematics but are not built for deep learning pipelines. TEMA relies heavily on Deep Neural Networks (DNNs) for extreme data analytics, such as the Concept-Based Content-Based Image Retrieval (CCBIR) pipeline. TEMA's custom architecture—leveraging Python-native distributed frameworks like Dask—allows for the processing of tensors, raster arrays, and neural network inferences within the exact same computational pipeline. Attempting to force these operations through a traditional SQL database creates severe I/O bottlenecks and data serialization overheads that are untenable during an emergency.

- **Temporal Awareness of Hazard Evolution:** A purely static layer, such as distance from the current fire perimeter, only answers the question of *current* proximity. In wildfire or flood operations, the more relevant question concerns how exposure evolves over the next time window. Existing tools typically map the "now." TEMA refines this by integrating simulation outputs (e.g., from FireSim or 3Di) as a temporal hazard source. Instead of generating a single distance-to-hazard indicator, the module produces time-indexed indicators derived from simulated front evolution (e.g., distance at multiple forecast steps, first threshold crossing time). This introduces the temporal variable directly into the geovisual layer, bridging current situational awareness with predictive response planning.

- **Mission-Oriented Heuristic Prioritization:** In emergency conditions, completeness of visualization is not sufficient; operators need relevance, ranking, and focus. The innovative layer in TEMA is the heuristic that sits between raw geospatial computation and final visual presentation. This heuristic transforms raw measurements into priorities by combining spatial proximity, infrastructure criticality (e.g., hospitals vs. empty sheds), expected event evolution, and configurable operational thresholds. The system is designed to foreground the subset of assets and areas that require attention first, rather than merely rendering all available layers.

- **Flexible Edge-to-Cloud Composability:** TEMA does not revolve around a single monolithic viewer or a centralized cloud database. Its geovisual capability is structured as a set of

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

reusable processing blocks packaged as independent Docker containers. Natural Disaster Management often occurs in environments with degraded network connectivity. TEMA's microservice architecture allows these analytical building blocks to be orchestrated via Kubernetes to run in a massive cloud cluster, or pushed to the "edge"—executing directly on a drone ground station or a local command vehicle's SmartDesk. This edge-to-cloud continuum ensures continuous analytical capability, a flexibility that monolithic GIS architectures struggle to replicate.

For these reasons, the innovation of TEMA should not be framed as the development of a generic viewer. Existing platforms remain highly valuable reference points and complementary technologies. However, the true added value of the TEMA Geovisual Analytics module lies in combining FIWARE-based event-driven orchestration, native AI processing, time-aware predictive analytics, and mission-oriented heuristics inside a coherent, edge-to-cloud emergency-management architecture.

4.4 Integration for the TEMA Visualization: Interfacing Geovisual Analytics with the Digital Twin, SmartDesk and Augmented Reality Visualization

Within Work Package 5, a central objective is the creation of a cohesive simulation and visualization ecosystem for emergency management. This is achieved through the tightly coupled interaction of three key tasks: T5.1 (Precise Digital Twin), T5.2 (Geovisual Analytics), and T5.3 (Smartdesk and XR Viewer). Task T5.2, Geovisual Analytics, serves as the critical processing and integration nexus, effectively bridging the foundational 3D environmental models from T5.1 with the advanced, interactive visualization capabilities developed in T5.3. Figure 2 depicts the integration workflow of the Geovisual Analytics with the other components.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

The workflow commences with T5.1, which is responsible for generating precise and detailed 3D models of the affected environment—the Digital Twins—often derived from sources such as drone imagery. These Digital Twin outputs provide the fundamental spatial context and geometric underpinnings for the subsequent analytical and visualization stages. The development in T5.1 has focused on establishing robust methods for creating these accurate 3D representations, which are then made available to the TEMA platform.

Task T5.2, Geovisual Analytics, then ingests these Digital Twin models from T5.1 alongside other relevant geospatial data streams originating from the broader TEMA platform, such as fused information products or outputs from geosocial media analysis. As an event-driven component, Geovisual Analytics is triggered by the arrival of new or updated data. It performs a range of crucial processing steps: complex spatial queries are executed (e.g., calculating distances from hazards like fires to various assets), the data is semantically enriched with contextual meaning, and it is filtered to highlight vulnerable or critical infrastructure elements. A cornerstone of this task is ensuring semantic interoperability across the platform. This is achieved through the rigorous application of FIWARE data models, which provide a standardized language for representing geospatial entities and their attributes. By transforming diverse inputs into these common models, Geovisual Analytics ensures that its outputs are consistently structured and readily interpretable by other TEMA components. These processed, enriched, and semantically interoperable data products are then published back, typically via the FIWARE Context Broker.

Finally, the components developed under Task T5.3—namely the Smartdesk and the XR Viewer—subscribe to and consume these refined geospatial data products from the FIWARE Context Broker, which have been prepared by T5.2. The Smartdesk leverages this data to present users with dynamic dashboards, 2D/2.5D map views displaying, for example, buildings highlighted according to risk levels, and tools for interacting with various data layers. Similarly, the XR Viewer utilizes these processed geospatial outputs to render 3D visualizations, allowing for a more intuitive understanding of the evolving situation. The efficiency of Geovisual Analytics in T5.2 directly contributes to the responsiveness of these T5.3 visualization tools, enabling timely presentation of critical information to end-users.

In essence, WP5 demonstrates a clear and functional progression of data and value. T5.1 establishes the detailed 3D environmental foundation. T5.2 acts as the intelligent intermediary,

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

transforming and enriching this foundational data along with other inputs, ensuring semantic consistency through FIWARE standards. T5.3 then delivers these actionable geospatial insights to the end-user through sophisticated and interactive visualization interfaces. Good progress has been made in all three tasks towards achieving their objectives and performance indicators related to accuracy, resolution, and responsiveness. The specific algorithms, detailed user interface features, and further advancements of the Precise Digital Twin (T5.1) and the visualization components like the Smartdesk and XR Viewer (T5.3) will be elaborated upon in subsequent project documentation focused on those specific areas, as the primary aim here is to delineate the crucial role of Geovisual Analytics (T5.2) in processing, integrating, and preparing data for effective visualization within WP5.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

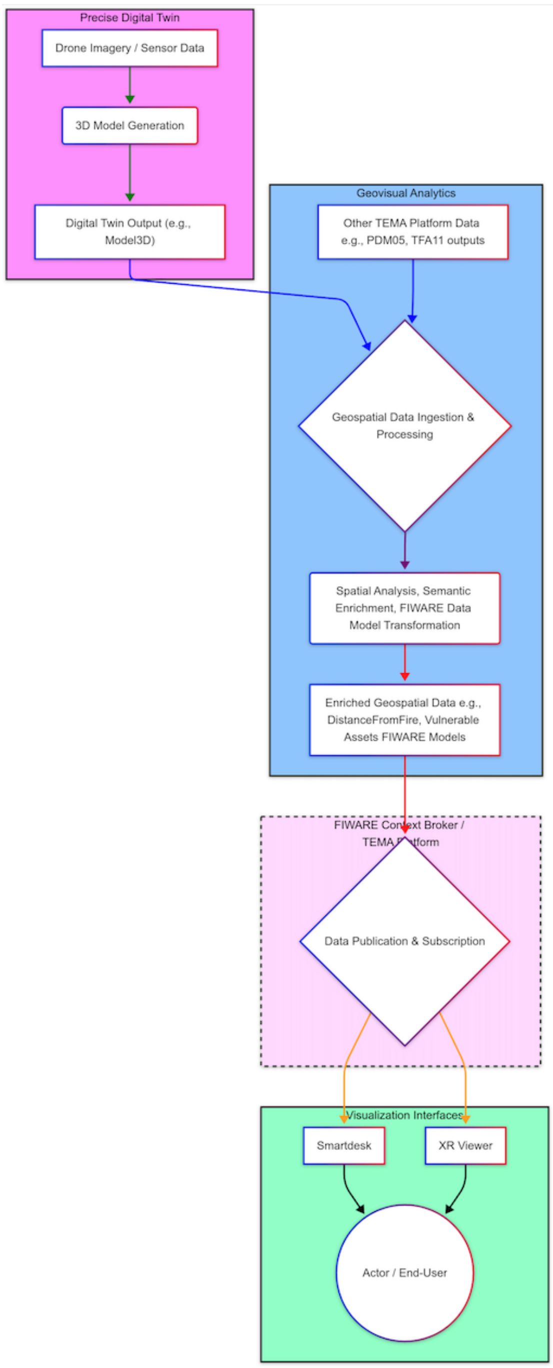


Figure 2. Integration workflow of the Geovisual Analytics with the Digital Twin, SmartDesk and Augmented Reality Visualization

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

4.5 Work in Geovisual Analytics in TEMA

The Geovisual Analytics component (Task T5.2) within TEMA focuses on transforming diverse, large-scale geospatial data into actionable intelligence for Natural Disaster Management (NDM), serving as a crucial link between data sources, the Digital Twin (T5.1), and visualization interfaces (T5.3). Key efforts center on developing modular "Analytics Building Blocks" for core geospatial operations like OpenStreetMap searches, GeoTIFF to COG conversion, and critical distance_from_fire/flood calculations. To handle "extreme data" volumes and achieve (near) real-time performance, these blocks heavily leverage Dask for efficient parallel processing and workload balancing across distributed environments like Kubernetes. This ensures scalability from municipal to regional disaster scopes. Seamless integration with the TEMA platform is achieved through the FIWARE Context Broker, adopting an event-driven architecture. Geovisual Analytics subscribes to new data entities (e.g., Maps4Fire), processes them, and publishes enriched, semantically interoperable results (e.g., DistanceFromFire) using standardized FIWARE data models. This allows components like the SmartDesk to dynamically visualize insights, such as color-coding buildings by fire risk based on distance, and supports alert generation. Continuous collaboration with visualization partners (e.g., KAMK) and end-users through hackathons drives iterative refinement of algorithms and outputs.

Visualizing the processed data is fundamentally important for the project because it allows non-technical operators to immediately manage the emergency. In **Figure 3**, we can see the fire-prone areas in red, while the coloured dots correspond to the buildings. Those tending towards red are those closest to the fire and, consequently, most at risk; conversely, those in green are furthest from the fire and, momentarily, safe. The result appears differently in the SmartDesk application (Figure 4).

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

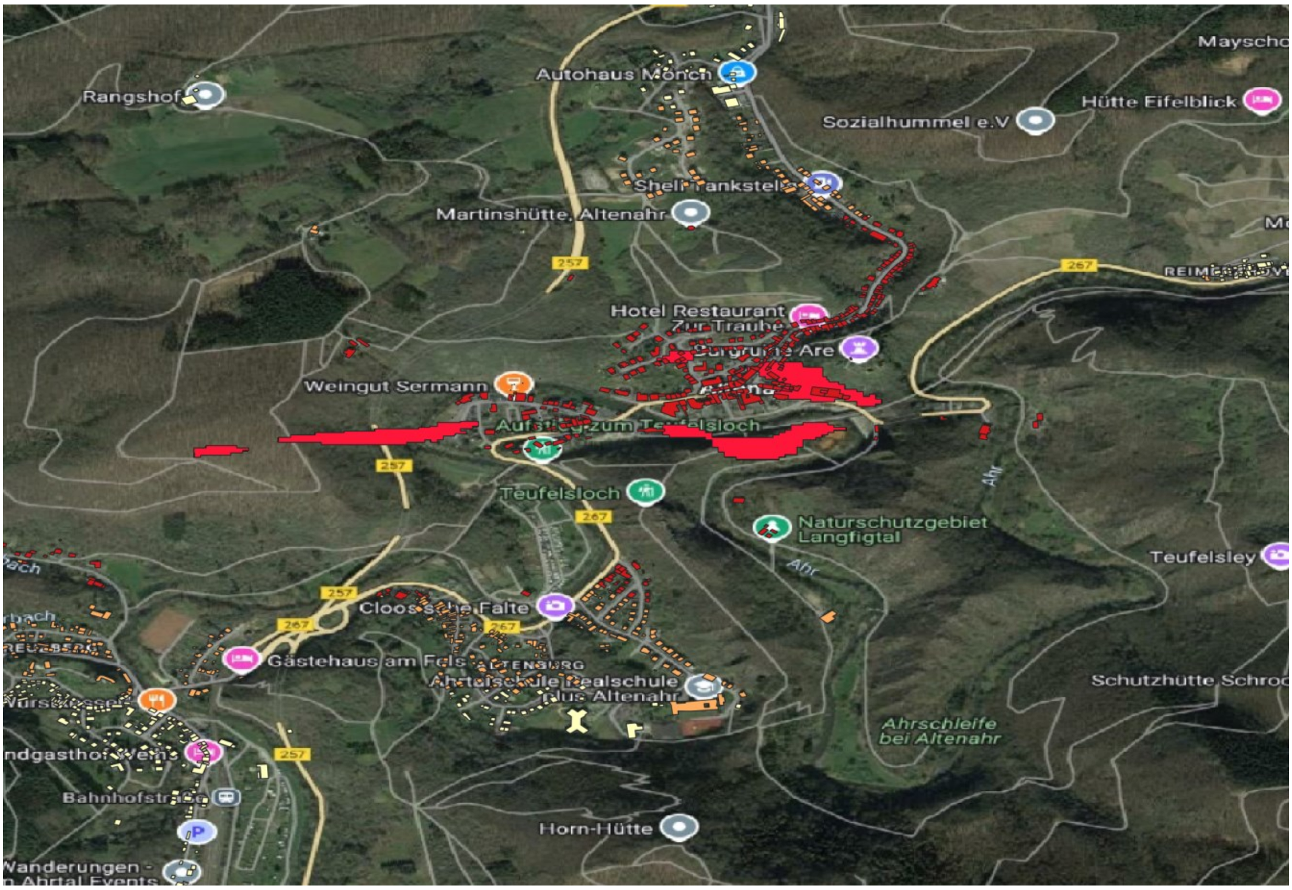


Figure 3. the result of geovisual analytics processing - fire case

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

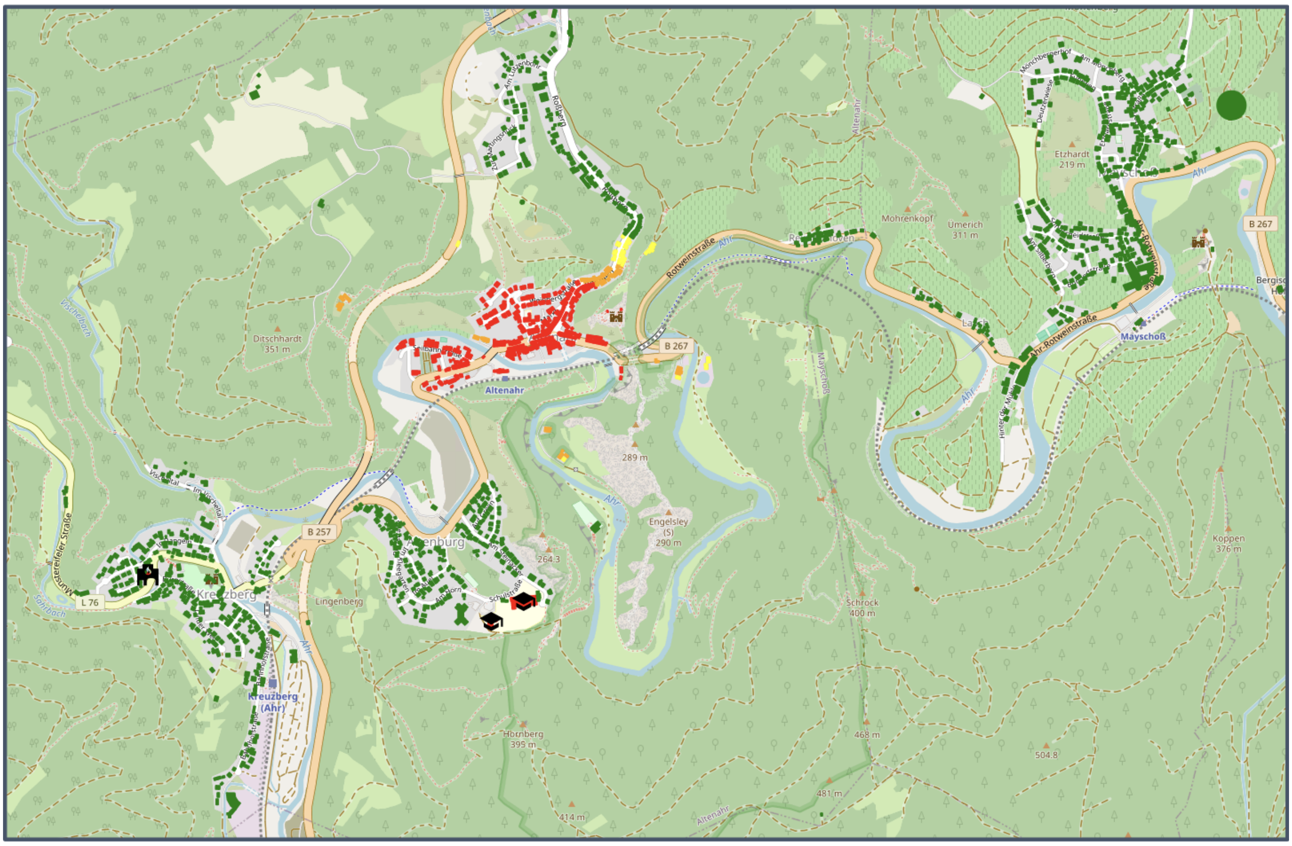


Figure 4. SmartDesk application: distance from fire

The Figure 4 illustrates the feature of the TEMA SmartDesk platform to dynamically display and filter buildings based on their real-time proximity to a detected fire, leveraging computations from the Geovisual Analytics module.

The map visualization depicts a geographical area with buildings. These buildings are color-coded (e.g., green for greater distances, transitioning through yellow and orange to represent increasing proximity to the fire), providing an immediate visual assessment of assets at varying risk levels based on their distance from the hazard. This visual representation directly reflects the output of the underlying geovisual analytic. The slider is featured empowers the user to dynamically adjust

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

the distance threshold for filtering buildings depending on their distance from fire. This allows for focused analysis on assets within specific user-defined proximity zones to the fire. This interactive filtering and visualization within the SmartDesk is powered by the underlying distance calculations continuously provided by the Geovisual Analytics module, demonstrating how processed insights are made directly actionable for the emergency manager. The system thereby provides an intuitive and responsive tool for operators to quickly assess risk, understand spatial relationships, and make more informed decisions during fire emergencies.

Figure 5. SmartDesk application: distance from fire real-time filtering

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

5. Core Components of the Geovisual Analytics Module

This chapter delves into the key components of geovisual analytics, exploring how they work together to create an integrated system for effective data analysis and visualization. It highlights the interconnections between these components in supporting decision-making processes.

5.1 Overview of the components

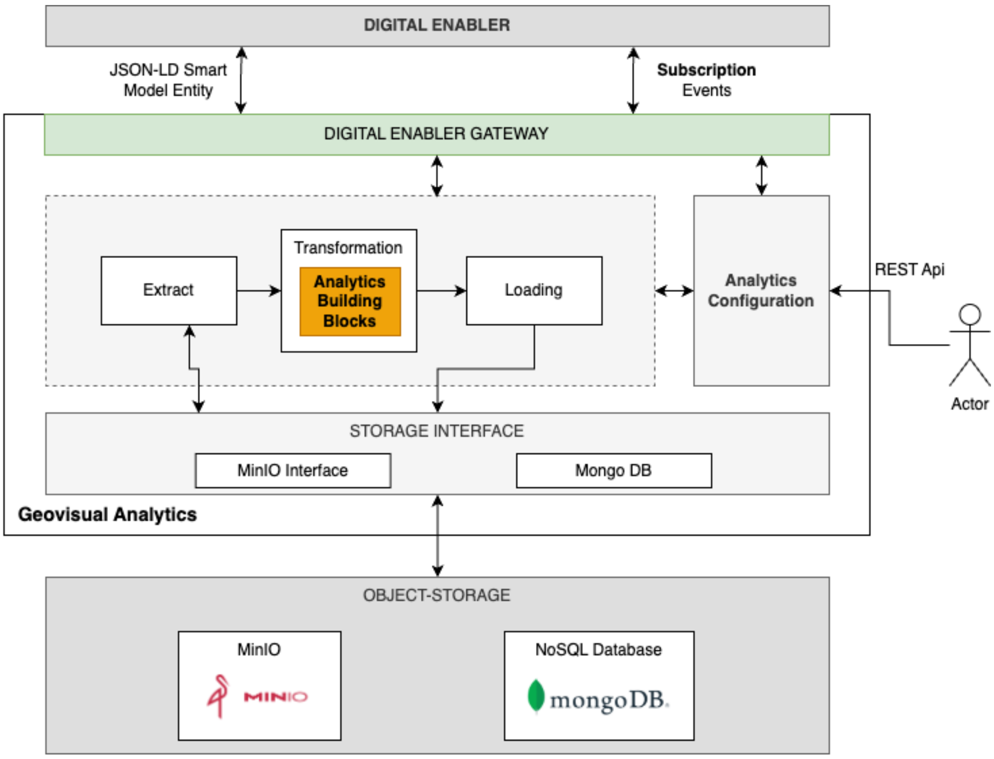


Figure 6. Geovisual Analytics, Architecture overview.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

The architecture of the geovisual analytics component follows a layered pattern that emphasises modularity, clear separation of concerns, and container-native deployment (Figure 6). The Geovisual Analytics architecture facilitates integration with the broader TEMA ecosystem, particularly the Digital Enabler, and provides interfaces for data storage, retrieval, and external configuration.

Each layer handles a specific responsibility, which helps keep the system organized and manageable. This design focuses on building the system from independent parts that work together but can be developed and maintained separately. Additionally, it is intended to run in container environments like Docker, allowing the system to be easily deployed, scaled, and managed across different computing platforms. Overall, this approach promotes flexibility, clarity in the system's structure, and ease of deployment.

At its core, the Geovisual Analytics module features an ETL (Extract, Transform, Load) pipeline. This pipeline is responsible for the systematic processing of geospatial information.

- The Extract phase retrieves data, primarily interacting with the underlying Storage Interface.
- The Transformation phase is central to the module's intelligence, utilizing a set of Analytics Building Blocks. These blocks encapsulate the specific algorithms and logic for performing advanced geovisual analyses, such as distance calculations, spatial queries, and semantic enrichment.
- The Loading phase then writes the processed and transformed data back, again through the Storage Interface.

The processing layer is where compute-intensive operations occur. It is built around `rasterio` for low-level raster input–output, `rioxarray` for labelled n-dimensional arrays, and `dask` for distributed task scheduling. Raster chunks are mapped onto `dask` workers inside the same pod when dealing with moderate data sizes, or across multiple pods when the cluster has spare

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

capacity. Vector operations reuse `pygeos`[1] and `shapely`[2], accelerated by spatial indexes, and can be executed in parallel when geometries exceed a configurable threshold. This hybrid strategy delivers high throughput without exhausting resources on smaller workloads.

Interaction with the Digital Enabler is managed via a Digital Enabler Gateway. The Geovisual Analytics module receives JSON-LD Smart Model Entities from the gateway, likely representing new or updated geospatial data requiring processing. It also responds to or is triggered by Subscription Events originating from the Digital Enabler, indicating an event-driven nature for initiating analytical workflows. This gateway exposes a single entrypoint for context-broker notifications and a complementary REST interface for ad-hoc analytic requests. The gateway authenticates incoming messages, validates their schema against predefined JSON documents, and places accepted jobs onto a lightweight in-memory queue. Immediate acknowledgement to the context broker ensures that upstream applications remain decoupled from processing latencies (Figure 7).

[1] https://pygeos.readthedocs.io/en/stable/

[2] https://shapely.readthedocs.io/en/stable/

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

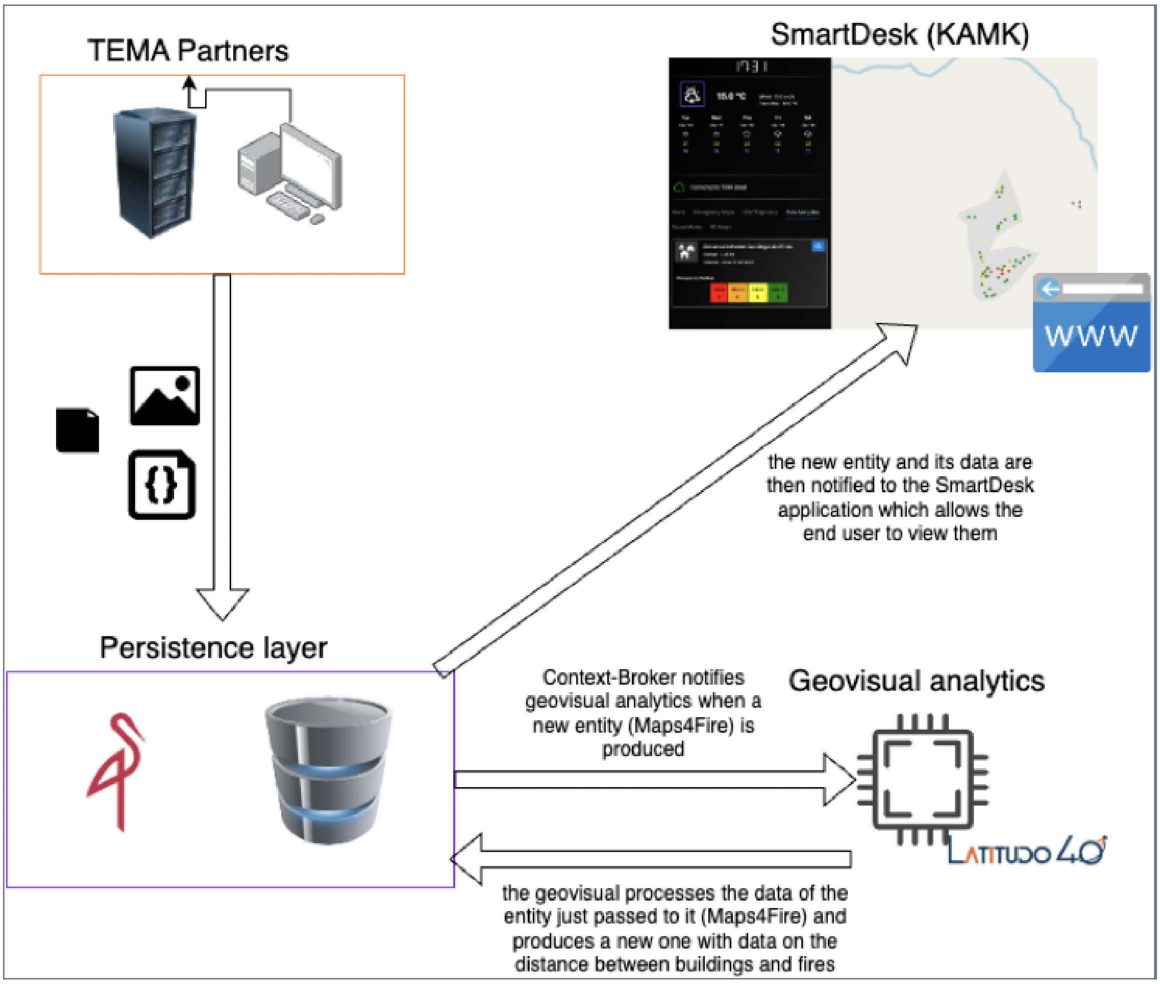


Figure 7. Interaction of the Geovisal Analytics components.

The Storage Interface layer within the Geovisual Analytics module provides an abstraction over various data storage technologies. It includes:

- A MinIO Interface for interacting with S3-compatible object storage.
- A Mongo DB interface, suggesting interaction with a NoSQL document database.

The persistence layer manages object storage interaction via the MinIO Python SDK. Artifacts such as COGs, GeoJSON files, and Parquet tables are uploaded with versioned keys that encode the job

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

identifier, the execution timestamp, and the checksum of the configuration file. The same layer maintains a metadata registry inside the context broker, where each digital object is represented by an NGSI-LD entity carrying spatial extent, resolution, coordinate-reference system, lineage, and minimal statistics.

External configuration and potentially ad-hoc analytical requests are managed through an Analytics Configuration block. This block is accessible to an Actor (e.g., a system administrator, developer, or another automated service) via a REST API, allowing for dynamic adjustments to the analytical processes and parameters.

In summary, the Geovisual Analytics architecture is characterized by:

- An event-driven ETL pipeline managed for data processing.
- A modular approach using "Analytics Building Blocks" for flexible transformation logic.
- A versatile storage interface supporting object storage (MinIO) and NoSQL databases (MongoDB).
- Clear integration points with the Digital Enabler for data input and event subscription.
- A REST API for external configuration of analytics, providing control and adaptability.

This design enables the Geovisual Analytics module to handle diverse geospatial data, perform complex analyses, and integrate seamlessly within the event-driven TEMA platform.

5.3 Data Processing and Workflow

The Geovisual Analytics module is designed to handle a diverse array of geospatial data sources and types, orchestrating their processing through a sophisticated ETL (Extract, Transform, Load) pipeline that leverages parallelization for efficient handling of large datasets. The module is capable of ingesting and processing a wide variety of data crucial for comprehensive situational awareness and emergency management. These include drone-derived imagery, satellite

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

imagery, meteorological data, social media data, and digital terrain models. Crucially, for vector data highlighting critical infrastructure and areas of interest (such as public halls, hospitals, and schools), the pipeline prioritizes the ingestion of high-quality, authoritative geospatial data from European and National Spatial Data Infrastructures (SDIs), Copernicus Emergency Management Service products, and local Civil Protection databases. Crowd-sourced data (e.g., OpenStreetMap) is ingested strictly as a supplementary layer to enrich non-critical contextual information or act as a fallback where official data might be temporarily unavailable.

The core of the data processing workflow is an ETL pipeline. This ensures a systematic and repeatable approach to data handling:

- Extract: In this initial phase, relevant data is retrieved from the TEMA platform's storage systems (e.g., MinIO for object storage, MongoDB for NoSQL data), directly from incoming event streams via the Digital Enabler Gateway, or queried from external authoritative data sources. This involves accessing data in its various native formats.
- Transform: This is the most computationally intensive phase and where the "Analytics Building Blocks" come into play. Raw data undergoes a series of transformations:
    - Harmonization and Standardization: Data from disparate sources is converted into common formats and coordinate reference systems.
    - Semantic Enrichment: Contextual information is added, and data is structured according to FIWARE data models to ensure semantic interoperability across the TEMA platform. During this phase, the pipeline applies a data hierarchy, prioritizing authoritative infrastructure geometries and augmenting them with supplementary open-source datasets (like OSM), only when necessary to complete the spatial context.
    - Analytical Operations: Specific geospatial analyses are performed based on the requirements of the active emergency management scenario. This can include spatial queries (e.g., point-in-polygon tests for identifying buildings within a flood zone), proximity calculations (e.g., distance of assets from a fire front), raster operations (e.g., deriving slope from a DTM), and combining different data layers (e.g., overlaying vector data of critical infrastructure on satellite imagery).
- Load: Once processed and transformed, the enriched data and analytical results are loaded back into the TEMA storage systems (MinIO, MongoDB). Importantly, metadata and

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

notifications about the newly available or updated data are published to the FIWARE Context Broker, making these insights accessible to other TEMA components like the SmartDesk and XR Viewer.

Handling the sheer volume and velocity of geospatial data in emergency scenarios necessitates efficient parallelization. The Geovisual Analytics module heavily relies on **Dask** to achieve this. Dask introduces genuine *big-data* capabilities into the pipeline. Dask is a flexible parallel computing library for analytics that integrates natively with Python and popular data science libraries like NumPy, Pandas, and Scikit-learn, which are often used within the "Analytics Building Blocks."

- **Task Parallelism for ETL**: Dask is employed to break down large ETL tasks into smaller, independent chunks that can be processed in parallel. For instance, when processing large raster datasets (e.g., satellite images), Dask can divide the raster into tiles. Each tile can then be processed by a separate Dask worker, whether on a single multi-core machine or distributed across a cluster. Similarly, operations on large collections of vector features can be parallelized.

- **Lazy Evaluation**: Dask often uses lazy evaluation, meaning computations are only executed when explicitly requested. This allows for the construction of complex computational graphs representing the entire transformation pipeline. Dask's scheduler then optimizes this graph for parallel execution, efficiently managing dependencies and data movement.

- **Distributed Computing**: For very large datasets that exceed the memory or processing capacity of a single machine, Dask can scale out to a cluster of machines (e.g., a Kubernetes cluster). It manages the distribution of tasks and data across the cluster, handling communication and data locality to minimize overhead.

- **Workload Balancing**: Dask's dynamic task scheduler inherently provides workload balancing by distributing tasks to available workers. This ensures that processing resources are utilized effectively and helps to avoid bottlenecks, leading to faster overall processing times.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

- **Out-of-Core Processing**: For datasets larger than available RAM, Dask supports out-of-core (or disk-based) computation, allowing operations on datasets that would otherwise be impossible to process in memory. This is particularly relevant for massive satellite imagery archives or high-resolution DTMs.

By leveraging Dask, the Geovisual Analytics module can efficiently process large and diverse geospatial datasets, performing complex transformations and analyses in a (near) real-time manner, which is critical for delivering timely insights during emergency management operations. The parallel execution capabilities ensure that the system can scale to meet the demands of various disaster scenarios, from localized incidents to large regional events.

By chunking multi-gigabyte rasters and partitioning millions of vector features across a dynamic worker pool, the system achieves near-linear speed-ups while maintaining a bounded memory footprint. Recent benchmarks show a drop from 40 s to 8 s when scaling distance-from-fire computations to eight workers, and a 300 GB Sentinel-2 mosaic converted to COG in 35 minutes on a four-node cluster. The tight integration between `dask`, `rasterio` and `rioxarray` lets prototype algorithms locally and then execute them unchanged at cluster scale, an innovation that accelerates both method development and operational deployment.

The Figure 8 shows how data is broken down in Dask, and then an orchestrator handles organizing which parts to solve first and how to reassemble them.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

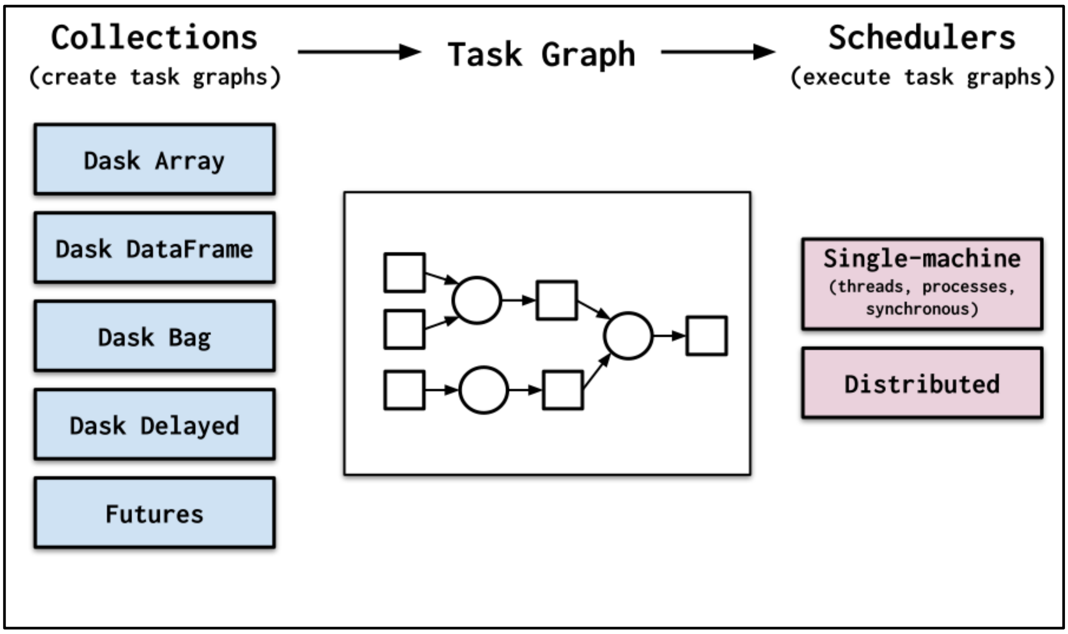


Figure 8. Dask working flow

This Figure 8 illustrates the core architecture and workflow of `Dask`, a parallel computing library in Python used for big data analytics, which is highly relevant for the geovisual analytics component.

- On the **left**, the figure shows the **Collections** layer, which includes high-level data abstractions such as `Dask Array`, `Dask DataFrame`, `Dask Bag`, `Dask Delayed`, and `Futures`. These collections allow users to build computational tasks in a familiar way (similar to NumPy arrays, pandas DataFrames, or Python lists), but without executing them immediately.

- In the **middle**, the **Task Graph** represents the abstraction of the computation itself. When you create operations on these collections, Dask constructs a directed acyclic graph (DAG) where nodes represent tasks and edges represent dependencies between them. This graph defines how tasks depend on and feed into each other.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

- On the **right**, the **Schedulers** execute the task graphs. There are two main types shown:
    - **Single-machine scheduler**: Runs tasks on a single computer, leveraging threads, processes, or synchronous execution.
    - **Distributed scheduler**: Runs tasks across multiple machines in a cluster, enabling true distributed parallel computation.

This architecture allows `Dask` to scale computations from a single machine to large clusters transparently. Users write code against familiar data collections, which are lazily translated into task graphs. These graphs are then scheduled and executed efficiently, either locally or distributed, enabling scalable and parallel processing of large datasets. This approach is fundamental in geovisual analytics for handling large raster and vector data in parallel, improving speed and resource utilization.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

5.4 Interface with TEMA components

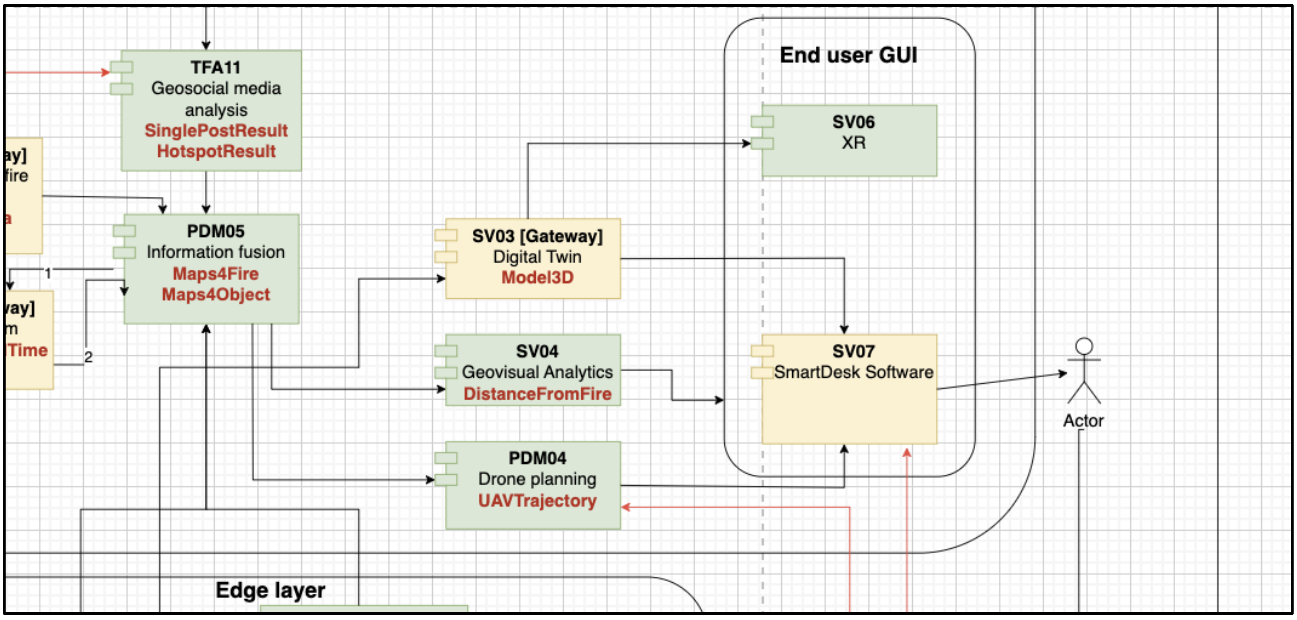


Figure 9. Component UML diagram - geovisual interaction with other components

Geovisual analytics occupies a pivotal position inside the TEMA ecosystem, acting as both consumer and producer of geospatial knowledge. As shown in Figure 9, its primary upstream relationship is with the digital enabler, which hosts the NGSI-LD context broker. All producers of spatial data, including drone-image pipelines, satellite down-link stations, and simulation services, publish metadata entities to the broker. Geovisual analytics subscribes to the entity types relevant to each business mission, for example DroneImage or Maps4Fire, and receives notifications whenever new datasets are available. This subscription mechanism eliminates polling delays and guarantees immediate reaction to evolving phenomena.

Once processing is complete, geovisual analytics updates or creates new entities, such as DistanceFromFire or DistanceFromFlood, in the same context broker. Downstream consumers consequently receive fresh notifications. These consumers include SmartDesk for two-dimensional dashboards, the XR-viewer for immersive three-dimensional exploration, and decision-support services that feed early-warning messages to first responders. By adhering to the same NGSI-LD

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

data model, the component ensures data lineage traceability across multiple hops, thereby fulfilling TEMA's accountability requirements.

Another critical relationship exists with object storage. The component relies on MinIO for durable binary storage. All intermediate raster tiles and final analytical products are stored in bucket hierarchies that mirror the logical grouping of business missions. Bucket names, object keys, and pre-signed download links are passed through the context broker so that consumer services can retrieve data without knowledge of internal storage credentials, satisfying security and least-privilege guidelines.

Geovisual analytics also interoperates with the precise digital-twin construction service led by Northdocks. The twin consumes cloud-optimised GeoTIFFs, 3D tiles, and high-resolution vector overlays generated by the analytics engine. In return, it produces derived geometries, such as extruded building footprints, that geovisual analytics can ingest as context layers for subsequent analyses.

The subsections below delve into the specifics of this interface.

5.4.1 Integration via FIWARE: Data Models and APIs

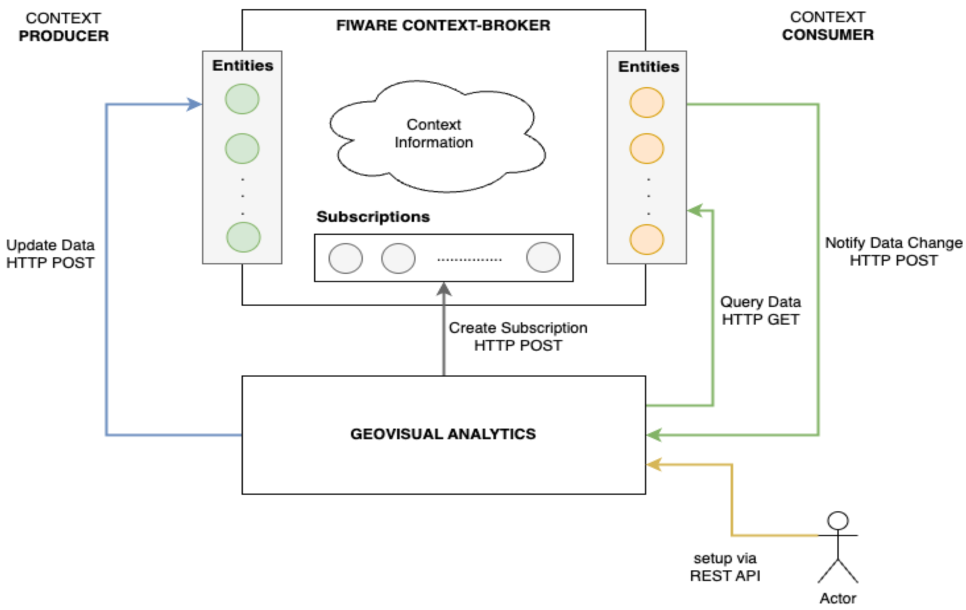


Figure 10. Interoperability through Fiware

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

The integration of the Geovisual Analytics module into the TEMA platform is achieved through a tightly coupled, event-driven architecture centered on the FIWARE Context Broker, which serves as the primary Digital Enabler (DE). This setup enables interoperability across platform components by adopting a decoupled publish/subscribe model, wherein Geovisual Analytics functions as both a consumer and producer of context information.

The data Flow and the interaction with the FIWARE Context Broker can be summarized in the following points:

- **Context Consumption via Subscriptions**: The Geovisual Analytics module is designed to react to changes and new data within the TEMA ecosystem. It achieves this by creating Subscriptions to specific Entities of interest managed by the FIWARE Context Broker. As shown in the diagram, Geovisual Analytics sends a "Create Subscription HTTP POST" request to the Context Broker. When a Context Producer (e.g., another TEMA analytical module, a sensor data ingestor, or the Digital Enabler Gateway processing raw inputs) updates relevant data via an "Update Data HTTP POST" to its managed Entities, the Context Broker, honoring the subscription, notifies the Geovisual Analytics module. This notification, often an event, triggers the appropriate Geovisual Analytics pipeline to process the new or updated information.
- **Context Production and Data Models:** Upon completion of its analytical tasks (e.g., calculating "DistanceFromFire", identifying vulnerable assets, enriching Digital Twin outputs), Geovisual Analytics publishes its results back to the TEMA platform. It does this by acting as a Context Producer, sending an "Update Data HTTP POST" request to the FIWARE Context Broker. This request creates new Entities or updates existing ones within the Broker's "Context Information" store. Crucially, these Entities adhere to standardized FIWARE data models. While not explicitly detailed in this specific diagram block, the interaction with the "JSON-LD Smart Model Entity" at the Digital Enabler Gateway level (as seen in the broader architecture) strongly implies the use of NGSI-LD compliant Smart Data Models [44]. These models ensure that the geospatial information, both input to and output from Geovisual Analytics, is structured consistently and is semantically rich. This semantic interoperability is key, allowing diverse TEMA components to understand and correctly interpret the data provided by Geovisual Analytics.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

- **Enabling Context Consumers**: Once Geovisual Analytics has published its processed and enriched Entities to the FIWARE Context Broker, other TEMA components, acting as Context Consumers (such as the SmartDesk, XR Viewer, or other decision-support tools), can access this information. They can either perform a "Query Data HTTP GET" request to retrieve specific entities or, like Geovisual Analytics, subscribe to relevant entities and receive automatic "Notify Data Change HTTP POST" notifications when Geovisual Analytics updates them.

All interactions within the platform are facilitated through the standardized FIWARE NGSI-LD API. This includes operations for creating and managing entities, subscriptions, queries, and notifications. The FIWARE-centric design provides several key advantages:

1. **Semantic Interoperability**: Linked-data typing unifies data from satellites, UAVs, and simulations under a common ontology.
2. **Event-Driven Responsiveness**: Real-time updates enable sub-second propagation of insights to operational dashboards.
3. **Open-Standard Reproducibility**: Compliance with NGSI-LD supports replication of experimental setups and integration of external tools without bespoke development.

This architecture elevates Geovisual Analytics from a standalone processor to a fully integrated, standards-compliant node within a broader open-data ecosystem.

So, TEMA components interact through JSON-formatted entities transmitted via the FIWARE Context Broker, which provides the backbone for data exchange and synchronization. Key entity types include:

- **DroneImage**: Represents data from drones, including images and metadata related to the location and time of capture.
- **Maps4Fire**: Contains information on fire-related maps, including perimeters, intensity, and propagation models.
- **DistanceFromFire**: Indicates the distances between specific resources or points of interest and the fire front, providing a key indicator of risk.
- **DistanceFromFlood**: Similarly, records the distances between sensitive areas and flooded zones during flood events, essential for emergency management.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

This communication system, based on standardized JSON entities and the FIWARE Context Broker, allows for seamless integration and immediate responsiveness, crucial elements for real-time emergency management operations.

DistanceFromFire

Table 2. DistanceFromFire json model for entity


```
{

        "id": "urn:ngsi-ld:LAT40:P:geovisualAnalyticsDistanceFromFire:01",
        "type": "geovisualAnalyticsDistanceFromFire",
        "description": {
          "type": "Property",
          "value": "Buildings at risk based on proximity to the natural event"
        },
        "bucket": {
          "type": "Property",
          "value": "lat40"
        },
        "creator": {
          "type": "Property",
          "value": "LAT40"
        },
        "file_name": {
          "type": "Property",
          "value": "01_distance_from_fire_20250520_150839.geojson"
        },
        "minio_url": {
          "type": "Property",
          "value":                                                      "https://storage.tema.digital-
enabler.eng.it/lat40/01_distance_from_fire_20250520_150839.geojson"
        },
        "resource_creation": {
```


This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

```
      "type": "Property",
      "value": "2025-05-20T15:08:41Z"
    },
    "location": {
      "type": "GeoProperty",
      "value": {
        "type": "Polygon",
        "coordinates": [
          [
            [
              29.0962433,
              64.3557818
            ],
            [
              29.1948516,
              64.3557818
            ],
            [
              29.1948516,
              64.4081359
            ],
            [
              29.0962433,
              64.4081359
            ],
            [
              29.0962433,
              64.3557818
            ]
          ]
        ]
      }
    }
```


This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

```
    }
```

Here (Table 2) is the description of each field in the JSON, presented in a Table 1.

Table 3. Description of DistanceFromFire entity

Field	Description
\`id\`	Unique identifier of the \`geovisualAnalyticsDistanceFromFire\` entity, in the format \`urn:ngsi-Id:LAT40:P:geovisualAnalyticsDistanceFromFire:01\`. This is a URI (Uniform Resource Identifier) specific to the NGSI-LD (Next Generation Service Interface - Linked Data) system used for data management and interoperability. It indicates that this entity is a distance from fire calculated by the LAT40 partner.
\`type\`	Type of the entity, in this case \`geovisualAnalyticsDistanceFromFire\`. This field specifies the nature or class of the data entity.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

`description`	Property that provides a textual description of the entity. The value `"Buildings at risk based on proximity to the natural event"` indicates that this data concerns buildings at risk based on their proximity to a natural event, presumably a fire. It includes a `type` field specifying `"Property"` and a `value` field containing the actual description.
`bucket`	Property that specifies the storage bucket where the data is saved, in this case `"lat40"`. It includes a `type` field specifying `"Property"` and a `value` field that identifies the specific bucket.
`creator`	Property that indicates the creator or organization that generated the data, in this case `"LAT40"`. It includes a `type` field specifying `"Property"` and a `value` field that identifies the author.
`file_name`	Property that specifies the name of the file containing the geographic data, in this case `"01_distance_from_fire_20250520_150839.geojson"`. The filename suggests that it is a GeoJSON file containing the distance from fire data, with a timestamp embedded in the filename. It includes a `type` field specifying `"Property"` and a `value` field that provides the filename.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

`minio_url`	Property that provides the URL to access the geographic data file in the MinIO storage, in this case `"[https://storage.tema.digital-enabler.eng.it/lat40/01_distance_from_fire_20250520_150839.geojson](https://storage.tema.digital-enabler.eng.it/lat40/01_distance_from_fire_20250520_150839.geojson)"`. This URL allows direct access to the file in the storage system. It includes a `type` field specifying `"Property"` and a `value` field that contains the URL.
`resource_creation`	Property that specifies the creation timestamp of the resource, in ISO 8601 UTC format, in this case `"2025-05-20T15:08:41Z"`. It indicates when the data was generated. It includes a `type` field specifying `"Property"` and a `value` field that contains the timestamp.
`location`	Property that defines the geographic location of the entity. The value is a GeoJSON object that describes a polygon. The polygon is defined by an array of coordinates that represent the vertices of the polygon. It includes a `type` field specifying `"GeoProperty"` and a `value` field that contains the GeoJSON object.

DistanceFromFlood

Table 4. DistanceFromFlood json model for entity

```
{
```


This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

```
  "id": "urn:ngsi-ld:LAT40:P:geovisualAnalyticsDistanceFromFlood:01",
  "type": "geovisualAnalyticsDistanceFromFlood",
  "description": {
    "type": "Property",
    "value": "Buildings at risk based on proximity to the natural event"
  },
  "bucket": {
    "type": "Property",
    "value": "lat40"
  },
  "creator": {
    "type": "Property",
    "value": "LAT40"
  },
  "file_name": {
    "type": "Property",
    "value": "01_distance_from_flood_20250416_130413.geojson"
  },
  "minio_url": {
    "type": "Property",
    "value":                                     "https://storage.tema.digital-
enabler.eng.it/lat40/01_distance_from_flood_20250416_130413.geojson"
  },
  "resource_creation": {
    "type": "Property",
    "value": "2025-04-16T13:04:16Z"
  },
  "location": {
    "type": "GeoProperty",
    "value": {
      "type": "Polygon",
      "coordinates": [
```


This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

```
          [
            [
              6.9445573,
              50.4894041
            ],
            [
              7.0272846,
              50.4894041
            ],
            [
              7.0272846,
              50.5433274
            ],
            [
              6.9445573,
              50.5433274
            ],
            [
              6.9445573,
              50.4894041
            ]
          ]
        ]
      }
    }
  }
```


This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

Table 5. Description of DistanceFromFlood entity

Field	Description
\`id\`	Unique identifier of the \`geovisualAnalyticsDistanceFromFlood\` entity, in the format \`urn:ngsi-ld:LAT40:P:geovisualAnalyticsDistanceFromFlood:01\`. This is a URI (Uniform Resource Identifier) specific to the NGSI-LD (Next Generation Service Interface - Linked Data) system used for data management and interoperability. It indicates that this entity is a distance from a flood calculated by the partner LAT40.
\`type\`	Type of the entity, in this case \`geovisualAnalyticsDistanceFromFlood\`. This field specifies the nature or class of the data entity.
\`description\`	Property that provides a textual description of the entity. The value \`"Buildings at risk based on proximity to the natural event"\` indicates that this data concerns buildings at risk based on their proximity to a natural event, presumably a flood. It includes a \`type\` field that specifies \`"Property"\` and a \`value\` field that contains the actual description.
\`bucket\`	Property that specifies the storage bucket where the data is saved, in this case \`"lat40"\`. It includes a \`type\` field that specifies \`"Property"\` and a \`value\` field that identifies the specific

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

	bucket.
`creator`	Property that indicates the creator or organization that generated the data, in this case `"LAT40"`. It includes a `type` field that specifies `"Property"` and a `value` field that identifies the author.
`file_name`	Property that specifies the name of the file containing the geographic data, in this case `"01_distance_from_flood_20250416_130413. geojson"`. The file name suggests that it is a GeoJSON file containing data on the distance from the flood, with a timestamp embedded in the file name. It includes a `type` field that specifies `"Property"` and a `value` field that provides the file name.
`minio_url`	Property that provides the URL to access the geographic data file in the MinIO storage, in this case `"[https://storage.tema.digital-enabler.eng.it/lat40/01_distance_from_flood_20250416_130413.geojson](https://storage.tema.digital-enabler.eng.it/lat40/01_distance_from_flood_20250416_130413.geojson)"`. This URL allows direct access to the file in the storage system. It includes a `type` field that specifies `"Property"` and a `value` field that contains the URL.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

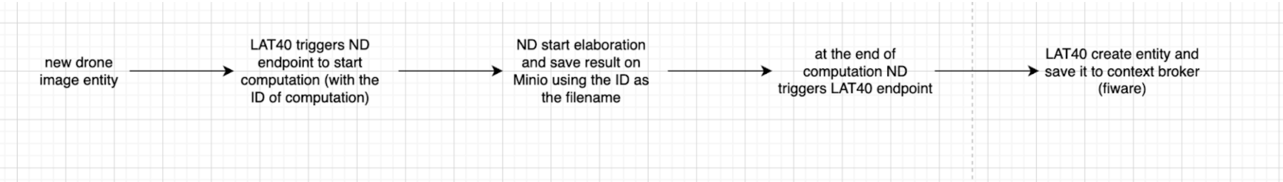
\`resource_creation\`	Property that specifies the creation timestamp of the resource, in ISO 8601 UTC format, in this case \`"2025-04-16T13:04:16Z"\`. It indicates when the data was generated. It includes a \`type\` field that specifies \`"Property"\` and a \`value\` field that contains the timestamp.
\`location\`	Property that defines the geographical location of the entity. The value is a GeoJSON object that describes a polygon. The polygon is defined by an array of coordinates representing the vertices of the polygon. It includes a \`type\` field that specifies \`"GeoProperty"\` and a \`value\` field that contains the GeoJSON object.

5.4.2 Integration with the Digital Twin

The work outlined in Deliverable 5.1 has progressed in the successful integration in the TEMA platform through a simple web API that has been developed and refined over the past several months. This represents a significant technological advancement in automated geospatial data processing and visualization as it does not require manual end-user input.

The newly developed RESTful API is specifically engineered to accept diverse image inputs, which are then automatically processed through advanced computational algorithms to generate high-fidelity 3D models for real-time use. Building strategically upon the foundational work established in Deliverable 5.1, which initially focused on developing core geospatial data processing capabilities and establishing robust visualization functionalities.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003



The processing pipeline employs state-of-the-art photogrammetry algorithms that transform raw image data into detailed 3D models specifically optimized for critical applications including real-time fire and flood risk assessment, comprehensive urban planning initiatives, emergency disaster response visualization, and advanced simulation environments.

The API infrastructure has been designed with a strong emphasis on secure and efficient data handling protocols receiving the data directly from the TEMA miIO server. By strategically focusing exclusively on image-based inputs, the API significantly simplifies the overall workflow for end-users while automating the traditionally complex and time-intensive process of 3D model creation. This approach maintains scalability and implements robust security measures to protect sensitive geospatial data throughout the processing pipeline.

Extensive collaborative efforts with key project partners from the TEMA consortium have ensured that the API comprehensively meets specific use case requirements across various operational scenarios. Regular testing sessions and iterative refinements have resulted in an simple but powerful API.

This integration achievement in the broader TEMA project initiative, enabling seamless access to automated 3D model generation capabilities for emergency responders. The effectiveness and reliability of these generated 3D models have been successfully demonstrated during the Finnish Historical trial, showcasing the system's capability to process historical imagery and reconstruct accurate 3D representations of past events. Additionally, the technology has been used and refined at multiple hackathons.

The generated 3D models can be efficiently displayed and manipulated through the SmartDesk interface, providing users with an intuitive platform for visualization, analysis, and decision-making. This integration ensures that complex 3D data remains accessible to users with varying levels of technical expertise.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

The API-driven integration architecture ensures exceptional scalability and maintains consistently low-latency performance characteristics, effectively supporting parallel processing capabilities for large-scale image datasets that may contain thousands of individual images. The system can handle concurrent processing requests while maintaining data integrity and processing quality standards. The collaborative refinement process has ensured that the API effectively meets TEMA's diverse use case requirements while advancing the fundamental goals established in Deliverable 5.1. This technological advancement enables fully automated, real-time 3D model generation capabilities and supports comprehensive geospatial analytics within the integrated TEMA platform ecosystem.

During the upcoming project months, our development will concentrate efforts on providing comprehensive support for the ongoing TEMA Trials, ensuring seamless integration and optimal performance across all testing scenarios. A primary focus will be placed on the continuous improvement and updating of the T5.1 algorithms to ensure they are fully optimized and specifically tailored to work effectively with both Flood and Fire Use Cases. Additional efforts will focus on expanding the API's capabilities to handle more diverse image formats and sources, while maintaining the high standards of accuracy and performance that have been established through the current implementation.

5.4.3 Integration with the Smart Desktop

The SmartDesk is a native desktop application, built using the Avalonia UI framework, and MapsUI open source library. The application works on both the Windows and Linux operating systems and under the hood it uses exclusively open source and cross-platform libraries, for example NetTopologySuite to process geo-spatial data and SharpKML to read and write KML files. An additional component from the KAMK side is used to connect SmartDesk and TEMA infrastructure. To provide a stable and real-time connection, SmartDesk uses the SignalR component.

This architecture allows the Smartdesk to receive updates from visual analytics component almost immediately. The users see a map layer with highlighted buildings and component card with additional information from the output, such as the amount of special buildings (eg. schools, hospitals, fire stations, police) (Figure 11). The component card displays a chart showing the total number of buildings, divided into four groups, based on their distance from the fire or water. By using the component card, the user can navigate to any area of interest, enable or disable the map layer, and control additional parameters such as distance from fire or water, or layer opacity.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

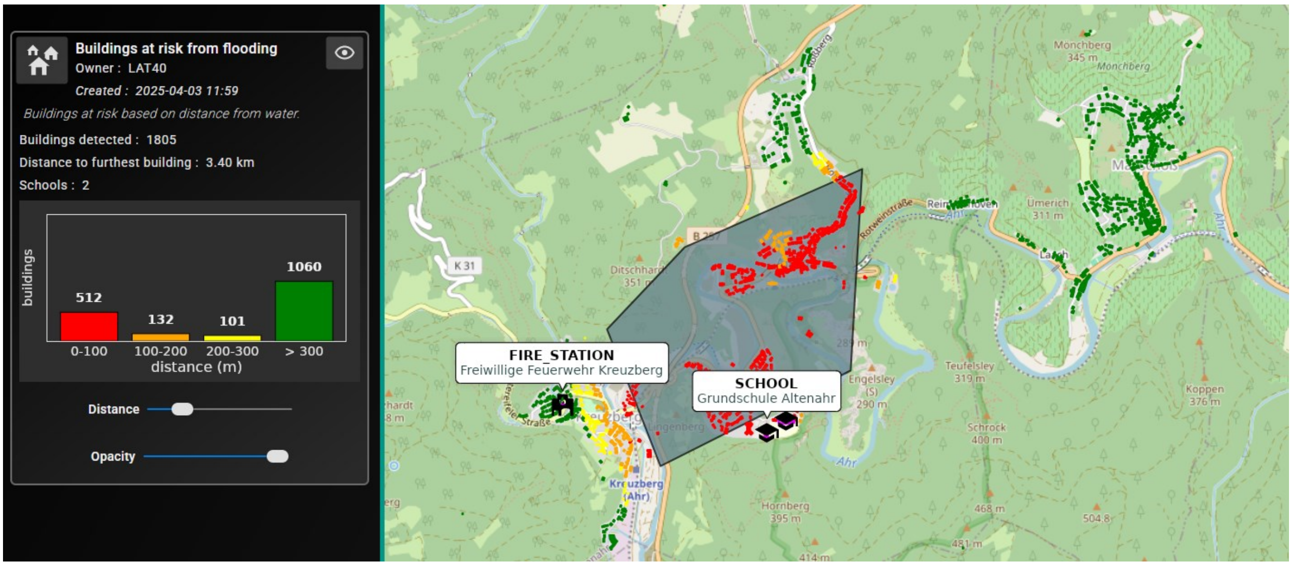


Figure 11. Visual analytics output from Ahrtal historical data

On the visualization front, the SmartDesk component developed by KAMK interacts with geovisual analytics through two channels. First, it subscribes to analytical result entities to update map layers dynamically. Second, it provides user feedback via a feedback API. When operators adjust visual-filter thresholds or request alternative colour ramps, SmartDesk emits parameter-tuning events that geovisual analytics interprets as configuration updates, thus closing the loop between analysis generation and human interpretation.

During hackathons the component collaborates with the alert-generation microservice. For instance, distance-to-fire metrics produced by geovisual analytics trigger alert entities when thresholds are exceeded. These alerts then propagate via push notifications to mobile devices used by first responders. Such cooperation demonstrates that the analytics engine not only enriches data but also initiates action across the broader platform.

Through these multifaceted relationships, geovisual analytics amplifies the value of every other subsystem in TEMA, turning isolated data streams into a coherent flow of situational intelligence.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

5.5 Geospatial functions

At the heart of geovisual analytics lies a library of atomic yet composable building blocks, each encapsulating a discrete geospatial operation. Building blocks follow a strict interface contract: each accepts a set of NGSI-LD entity identifiers or direct object-storage URLs, a JSON configuration payload specifying processing parameters, and returns a reference to the artifact produced along with a metadata entity describing it. This contract permits any block to be chained with any other, producing arbitrarily complex analytical pipelines while retaining clarity and testability.

5.5.1 Baseline Geospatial Data Strategy

To build an operational Digital Twin and perform mission-critical geovisual analytics, the quality and provenance of the underlying geospatial data are paramount. While crowd-sourced platforms like OpenStreetMap (OSM) offer exceptional global coverage and convenience—particularly during rapid prototyping and initial deployment—they cannot serve as the primary foundational layer when higher-quality, authoritative datasets exist. This is especially true in the European context, where national, regional, and municipal authorities maintain highly accurate, rigorously curated topographic, cadastral, and critical infrastructure datasets.

To satisfy this operational requirement while maintaining platform resilience, TEMA adopts a pragmatic, hierarchical data-source strategy, as illustrated in the figure below. The ingestion pipeline evaluates and retrieves data based on a strict priority gradient:

1. **Authoritative Local Datasets (Highest Priority):** The preferred option is the use of authoritative datasets provided by public authorities, Spatial Data Infrastructures (SDIs), or trusted operators for the specific area of interest (e.g., via WFS APIs or official regional portals).

2. **OpenStreetMap (OSM):** OSM serves as the primary, highly reliable fallback option. It is utilized strictly when authoritative data is temporarily unavailable, cannot be accessed under the required conditions during an active emergency, or lacks coverage for specific non-critical contextual features.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

3. **Overture Maps:** As an emerging, quality-checked open map dataset backed by major technology consortiums, Overture provides an additional layer of curated open data to supplement gaps.

4. **Other Sources:** Project-specific validated layers, drone-derived mapping, or commercial satellite annotations prepared for a given emergency trial.

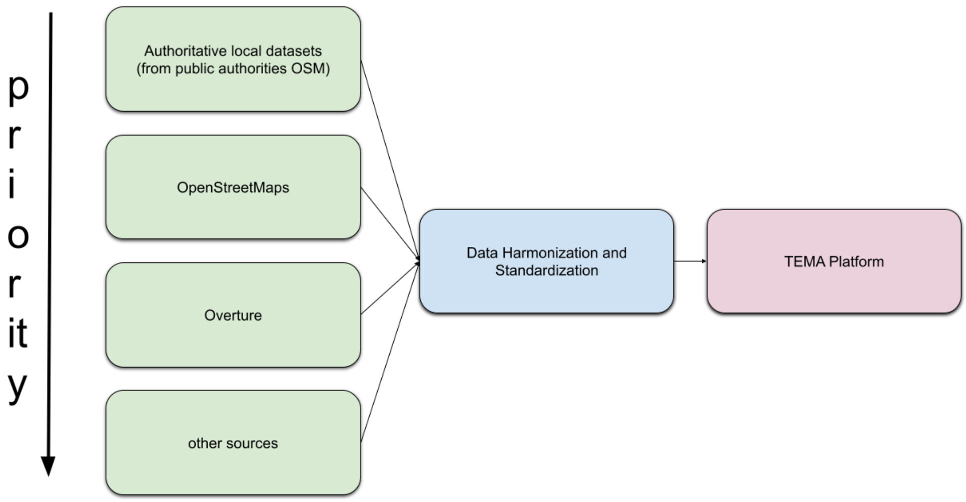


Figure 12 - Hierarchical Data Integration into TEMA platform

Crucially, as depicted in the workflow, all ingested data—regardless of its priority level—must pass through a **Data Harmonization and Standardization** layer before it is made available to the TEMA Platform. This step is a practical necessity. A single, automated, Europe-wide integration of authoritative data is currently unfeasible because local datasets vary significantly across borders in terms of schema, spatial resolution, update cycles, semantic classifications, and delivery formats.

Rather than attempting a brittle "one-size-fits-all" ingestion procedure, TEMA's harmonization layer evaluates the available data on a case-by-case basis. It resolves schema discrepancies and standardizes the diverse inputs into the unified FIWARE/NGSI-LD data models required by the downstream analytical blocks.

This architectural choice ensures that TEMA remains highly portable and resilient. For current trial areas, authoritative baseline data is prioritized and harmonized. In the future product evolution of

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

the tool, this same architecture provides a plug-and-play mechanism: municipalities, civil protection agencies, or infrastructure operators can securely connect their own validated baseline data via the harmonization layer. This seamlessly overrides the OSM/Overture fallbacks without requiring any alterations to the downstream geovisual analytics workflow, directly addressing the critical need for high-quality, authoritative geospatial sources in operational emergency management.

5.5.2 Analytics Building Blocks

Reflecting the data strategy outlined above, a foundational building block in the library is the **Contextual POI Ingestion block**. By combining spatial predicates with attribute filters, it retrieves points of interest such as hospitals, bridges, and fire stations within an area of interest. This block implements the hierarchical retrieval strategy: it prioritizes querying authoritative European and National Spatial Data Infrastructures (SDIs) via standardized APIs (e.g., WFS) for the trial regions, utilizing OpenStreetMap APIs strictly as a supplementary layer to enrich non-critical context or as a fallback. It outputs a unified GeoJSON feature collection stored in object storage, normalizing all retrieved geometries to a common coordinate reference system.

The **conversion block** transforms GeoTIFF rasters into cloud-optimised GeoTIFFs. Although apparently simple, the block performs tile re-ordering, internal overviews creation, compression adjustment, and adds spatial metadata such as bounding boxes and ground-control points to satisfy the COG specification. Because it operates on potentially multi-gigabyte files, the block leverages rasterio's windowed reading in combination with `dask` arrays, ensuring that only the necessary chunks reside in memory at any given moment.

The **distance block** calculates the nearest-neighbour distance between two vector layers. Internally it builds a spatial index over destination features using R-trees[3], then iterates over origin features, computing geodesic or planar distances depending on the projection. Optionally it filters origin features by attribute conditions, for example selecting only residential buildings with

[3] https://it.wikipedia.org/wiki/R-tree

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

occupancy greater than fifty. Distances are written back as an attribute to the origin layer, producing a new GeoJSON file ready for visualisation.

A **masking block** creates binary or multi-class masks from raster layers by applying threshold logic. Typical usage includes extracting burn scars from satellite imagery or isolating flood-affected pixels above a certain water depth. Masks become inputs to subsequent blocks, ensuring that downstream statistics ignore irrelevant pixels.

Every block publishes a provenance trail. The trail lists inputs, outputs, code version, execution duration, and resource usage. This information is emitted as an NGSI-LD entity, then harvested by TEMA's KPI dashboard to track how analytics performance improves over time. Provenance also allows reproducibility, because any analytical result can be regenerated by replaying the same sequence with identical parameters.

Testing frameworks accompany every block. Unit tests validate geometry handling, projection integrity, and edge-case behaviour. For example the zonal statistics block includes tests where raster and vector files have mismatched projections and verifies that automatic reprojection yields correct outputs. Load tests stress blocks with synthetic rasters of increasing size until performance metrics exit acceptable ranges, thereby guiding optimisation efforts.

The `distance_from_fire` building block quantifies how far critical assets lie from an active wildfire perimeter, enriching TEMA's common operating picture with a continuously updated proximity metric. The block consumes two primary inputs: an NGSI-LD entity representing the fire geometry, typically a real-time polygon or multi-polygon generated by remote-sensing components, and a vector layer containing the assets of interest, such as residential buildings, bridges, hospitals, or power lines. Both inputs can be supplied either as direct object-storage URLs or by reference to their entity identifiers, in which case the block retrieves the corresponding GeoJSON or GeoPackage files from the MinIO buckets declared in the metadata.

Internally the workflow begins by harmonising coordinate reference systems. If fire and asset layers use different CRSs, the block leverages `rioxarray` and `pyproj` to reproject assets onto the fire layer's datum, thereby preserving geodesic accuracy. A spatial index is then built over the asset geometries with pygeos R-trees, reducing the search space for nearest-neighbour queries. When the asset layer exceeds a configurable threshold—by default ten thousand features—the workload is partitioned automatically into balanced chunks and dispatched to a local or cluster-wide `dask`

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

scheduler. Each task computes the minimum distance between the chunk's geometries and the unioned fire perimeter using shapely.distance, returning results as NumPy arrays that are concatenated in constant time.

The output is a GeoJSON feature collection identical to the original asset layer but augmented with two new attributes: `distance_fire_m`, the shortest distance expressed in metres, and `risk_class`, a categorical field whose values low, medium, or high correspond to threshold intervals supplied in the block configuration. These thresholds are derived from response-time models validated by civil-protection partners during TEMA hackathons but remain fully user-customisable. The enriched layer is written back to object storage as a versioned key and is accompanied by a lightweight Parquet summary that aggregates asset counts per risk class, enabling rapid dashboard rendering without reloading the full geometry set.

Upon successful completion the block creates or updates an NGSI-LD entity of type `DistanceFromFire`. The entity stores the object key of the GeoJSON output, the time of computation, the software version, and simple statistics such as the minimum, maximum, and mean distance. This entity is then published to the Context Broker, triggering real-time updates in SmartDesk and alert-generation microservices. For example, when the high-risk count surpasses a predefined limit, SmartDesk automatically highlights the affected buildings in red and emits a notification card for incident commanders.

Configurability is exposed through a concise JSON schema. Optional parameters allow filtering the asset layer by attribute expressions, for instance processing only buildings where `usage` equals `school`, or clipping the analysis to a buffered envelope around the fire perimeter to save resources. Because the building block follows the common interface contract, it can be inserted into any analytical chain. Typical pipelines run a Critical Infrastructure Retrieval query search to collect critical infrastructure, feed the results into `distance_from_fire`, then pass the enriched layer to the zonal statistics block for population-exposure estimates.

From a DevOps perspective the block is packaged in its own Docker image, inheriting the base geovisual analytics image and adding only the libraries required for spatial indexing. Continuous-integration tests validate numerical correctness against synthetic datasets where analytic distances are known a priori, while stress tests ensure that performance targets are met under peak loads.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

In summary, `distance_from_fire` converts raw wildfire perimeters and static asset inventories into actionable, risk-aware insights. By delivering precise distance metrics, classifying assets according to configurable thresholds, and publishing results through standard TEMA channels, the block empowers emergency managers to prioritise inspections, allocate resources, and issue timely evacuation orders, thereby fulfilling a critical requirement of the fire business mission.

Time-aware wildfire exposure analytics

The current distance_from_fire building block can be further refined by consuming not only the observed fire perimeter at the current time, but also forecast perimeters generated by FireSim for successive time horizons. In this configuration, the building block computes a time series of asset-to-fire distances rather than a single static proximity value. Derived indicators may include minimum predicted distance over the simulation horizon, first time of entry into a given risk threshold, persistence of exposure within the high-risk class, and expected trend of hazard approach or recession. The output is therefore transformed from a static geospatial layer into a spatio-temporal analytic product that can be directly consumed by SmartDesk, the Digital Twin, and XR visualisation components.

This extension is particularly important in operational wildfire response because the current distance from the fire front does not fully capture the expected short-term evolution of risk. Two assets with identical present-time distance may have very different operational priority if one is located along a predicted propagation corridor while the other is not. By integrating FireSim outputs into the geovisual analytics workflow, TEMA introduces the temporal variable into the exposure assessment and enables more decision-oriented products such as time-aware risk classes, alert prioritisation, and timeline-based visual exploration.

5.6 Geospatial Image Retrieval

Content Based Image Retrieval (CBIR) is the process of retrieving similar images from a large database, given a query image, by analyzing and extracting their fundamental visual characteristics [32]. Recent advancements in Deep Learning have significantly enhanced CBIR, leveraging models capable of extracting meaningful representations from images. Despite these advancements, CBIR

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

for Natural Disaster Management (NDM) scenarios remains an underexplored area, with limited datasets specifically tailored for emergency response applications. Existing approaches, such as NetVLAD [33] and MixVPR [34], primarily focus on place recognition and generic retrieval tasks, which may not align perfectly with the needs of disaster response.

AUTH evaluated several state-of-the-art CBIR architectures, including NetVLAD [33], PatchNetVLAD [37], MIXVPR [34], CosPlace [35], EigenPlaces [36 and AnyLoc [38] for their ability to retrieve relevant images in disaster-related datasets. The goal was to identify a model that could effectively combine computational image representations and human semantic understanding while maintaining efficiency in large-scale CBIR tasks. Among the evaluated CBIR models, the proposed Concept-Based CBIR (CCBIR) pipeline, leveraging Non-Negative Matrix Factorization (NMF), demonstrated superior interpretability and CBIR accuracy compared to conventional methods.

To overcome the challenges of concept-based image retrieval in NDM scenarios, AUTH has evaluated its CCBIR pipeline on diverse images from three primary sources: 1) the European Flood 2013 dataset (251 query images, 100,192 reference images), evaluated for class-based retrieval tasks, 2) the PrePost Burnt Locations dataset (101 query images, 101 reference images), and 3) the Blaze [SIAVRAKAS 2024] wildfire dataset (1,100 query images, 4,308 reference images). These datasets encompass various NDM scenarios, including flood, wildfire, and post-disaster damage assessment imagery.

AUTH CCBIR pipeline extracts Deep Features using a task-specific DNN backbone, followed by NMF decomposition to derive semantic concepts and FAISS library for efficient CBIR. These concepts serve as structured representations that enhance retrieval performance by aligning image features with human-interpretable elements. Unlike traditional feature-based CBIR methods, AUTH's approach ensures that retrieved images are both visually similar and semantically relevant. An example of the CCBIR pipeline evaluated on the PrePost Burnt Locations Dataset is depicted on **Figure 12**.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

Figure 13. Result of the CCBIR pipeline in the PrePost Burnt Locations Dataset. The image on the left is the query, and the image on the right is the retrieved pre-disaster match.

AUTH CCBIR pipeline outperforms the baselines ones and scores 80.28% NDCG@100[4] on the European Flood 2013 dataset (~25.68% increase in performance), 0.6535 Recall@10[5] on the PrePost Burnt Locations dataset and 0.9776 Recall@5[6] on the Blaze dataset. Performance results are presented in detail in Table 5 and 6. The inference speed (s) was measured using a Nvidia GTX commercial GPU. In terms of speed, the CCBIR pipeline maintained competitive efficiency. On the PrePost Burnt Locations dataset, it had a total speed of 5.09 seconds, compared to MixVPR's 2.09 seconds but significantly outperformed Patch-NetVLAD's 58.64 seconds. On the Blaze dataset, it achieved a total speed of 152.12 seconds, closely matching MixVPR's 156.76 seconds, while remaining much faster than Patch-NetVLAD's 766.36 seconds.

[4] This metric evaluates the quality of ranked retrieval results, focusing on the top 100 items returned by a system. It accounts for both the relevance of retrieved items and their positions in the ranking, giving higher weight to relevant items appearing near the top. A value of 80.28% means the system's ranking is highly effective at placing relevant results early, with a 25.68% improvement compared to a baseline.

[5] This measures the proportion of relevant items found within the top 10 retrieved results. For example, a Recall@10 of 0.6535 means that about 65.35% of all relevant items are successfully retrieved among the first 10 results. It reflects the system's ability to find relevant matches quickly.

[6] Similar to Recall@10 but more stringent, Recall@5 measures the fraction of relevant items found within the top 5 results. A value of 0.9776 indicates that 97.76% of relevant items are retrieved in the first five results, demonstrating very high precision in the highest ranks.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

Table 6. Performance and Efficiency of CCBIR Methods on Blaze Dataset on Nvidia GTX 1080.

Method	Feature Extractor	Positives/Query	Recall@5	Recall@10	Speed(s)
Patch-NetVlad	VGG16	10	0.9615	0.9808	766.36
NetVlad	VGG16	10	0.9615	0.9744	112.31
EigenPlaces	ResNet50	10	0.9679	0.9808	149.30
CosPlace	ResNet50	10	0.9679	0.9744	150.34
AnyLoc	DINOv2-ViT-g/14	10	0.9744	**0.9872**	312.62
MIX VPR	ResNet50	10	0.9679	0.9744	156.76
CCBIR	ResNet18	10	**0.9776**	**0.9872**	152.12

Table 7. Performance and Efficiency of CCBIR Methods on PrePost Burnt Location Dataset on Nvidia GTX 1080.

Method	Feature Extractor	Positives/Query	Recall@5	Recall@10	Speed(s)
Patch-NetVlad	VGG16	1	0.2178	0.3663	58.64
NetVlad	VGG16	1	0.4356	0.5743	5.65
EigenPlaces	ResNet50	1	**0.5743**	0.6832	8.01
CosPlace	ResNet50	1	0.4356	0.6337	8.31
AnyLoc	DINOv2-ViT-g/14	1	0.5446	**0.7228**	354.18
MIX VPR	ResNet50	1	0.4950	0.6040	2.09
CCBIR	ResNet18	1	0.4554	0.6535	5.09

These results highlight the effectiveness of AUTH's CCBIR pipeline in balancing retrieval accuracy, interpretability, and efficiency, making it a robust solution for disaster-related CBIR tasks.

Comparison of visual place recognition methods for UAV imagery SOTA

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

Visual Place Recognition (VPR) is a critical task for UAV-based applications, enabling the retrieval of geo-referenced images for localization, disaster assessment, and infrastructure monitoring. In scenarios where GPS signals are lost or in GPS-denied environments, VPR serves as a key solution for UAVs to recognize previously visited locations. Traditional VPR approaches, such as NetVLAD [33] and Patch-NetVLAD [37], rely on Content Based Image Retrieval (CBIR) strategies where a query image is matched against a database of reference images based on visual similarity. More recent methods, such as MixVPR [34], CosPlace [35], EigenPlaces [36], and Anyloc [38], enhance feature representation learning through Deep Neural Networks (DNNs) and global/local descriptor aggregation. However, these methods have primarily been developed for street-level or on-road VPR applications, leaving UAV-based VPR as an underexplored research area.

To evaluate VPR performance for UAV imagery, AUTH conducted a comparative analysis of state-of-the-art methods across multiple UAV datasets, including ALTO [41], GL3D [43], VisLoc [42], and a custom Retrieval-Burnt Dataset, which features imagery captured before and after a wildfire. The evaluation focused on assessing CBIR accuracy across increasing database sizes to determine how well each method generalizes when the search space scales up.

Figure 14. On each column there is a representative query and reference image pair for each of the introduced datasets. The pairs belong to Burnt, VisLoc, and GL3D datasets from left to right

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

AUTH benchmark evaluation demonstrated that while pretrained VPR models perform well on multi-view datasets like GL3D and VisLoc, they struggle with multi-temporal settings where drastic environmental changes occur. Notably, Anyloc consistently outperformed other methods in domain generalization due to its foundation model-based feature aggregation. In contrast, Patch-NetVLAD and NetVLAD suffered significant performance degradation when database sizes exceeded 100k images, highlighting their sensitivity to large-scale CBIR challenges.

Table 5 presents the CBIR performance of VPR methods on the Burnt Dataset for Database Size = 10k images, showing how CBIR accuracy increases while the number of positive images per query increases. Similarly, Tables 6 and 7 illustrate the performance comparison of the 6 different SOTA methods for Database Size 10k and 200k images. Also, latency performance is depicted on Table 8, measuring feature extraction speed on an NVIDIA GTX 1080 Ti GPU.

Table 8. Performance Comparison on Burnt Dataset for Database Size = 10k.

Method	Recall@1	Recall@5	Recall@10	Recall@50	Recall@100
NetVlad	0.12	0.35	0.47	0.67	0.72
Patch-NetVlad	0.10	0.29	0.38	0.59	0.72
MIX VPR	0.19	0.47	0.51	0.65	0.70
CosPlace	0.05	0.17	0.23	0.31	0.37
EigenPlaces	0.03	0.25	0.33	0.41	0.43
AnyLoc	0.05	0.27	0.43	0.59	0.64

Table 9. Performance Comparison on ALTO Dataset for Database Size = 10k.

Method	Recall@1	Recall@5	Recall@10	Recall@50	Recall@100
NetVlad	0.06	0.18	0.29	0.73	0.85
Patch-NetVlad	0.20	0.56	0.72	0.84	0.85
MIX VPR	0.32	0.72	0.89	1.00	1.00
CosPlace	0.21	0.51	0.64	0.79	0.83

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

EigenPlaces	0.29	0.69	0.77	0.91	0.94
AnyLoc	0.35	0.81	0.94	1.00	1.00

Table 10. Performance Comparison on ALTO Dataset for Database Size = 200k.

Method	Recall@1	Recall@5	Recall@10	Recall@50	Recall@100
NetVlad	0.06	0.18	0.29	0.72	0.83
Patch-NetVlad	0.19	0.57	0.70	0.82	0.83
MIX VPR	0.42	0.82	0.93	0.98	0.99
CosPlace	0.21	0.51	0.63	0.75	0.77
EigenPlaces	0.28	0.62	0.73	0.86	0.90
AnyLoc	0.33	0.78	0.92	1.00	1.00

Table 11. Feature Space Search Speed on Nvidia GTX 1080.

Method	Feature Extraction Speed (s)
NetVlad	6.82
Patch-NetVlad	512.30
MIX VPR	14.12
CosPlace	6.64
EigenPlaces	4.70
AnyLoc	2.37

The results show that Anyloc achieves the best generalization, especially in multi-view UAV imagery, while MixVPR also performs well in terms of CBIR accuracy. However, multi-temporal settings remain challenging, with all methods struggling in the Burnt Dataset. Patch-NetVLAD and NetVLAD degrade significantly with increasing database size, highlighting scalability issues. Latency analysis confirms Anyloc as the most efficient, making it the best choice for real-time UAV-based VPR.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

Distilling Structural Knowledge for Class-based Image Retrieval

In the context of Natural Disaster Management (NDM), rapid and efficient image retrieval is essential for enabling real-time decision-making during critical response scenarios. However, state-of-the-art DNNs, while powerful, are often computationally expensive and unsuitable for deployment in low-resource or time-sensitive environments. To address this, we leverage knowledge distillation—a model compression technique that transfers knowledge from a large, high-performing teacher model to a lightweight student model. By optimizing the student to mimic the representational structure of the teacher, we enable efficient Content-Based Image Retrieval (CBIR) while reducing computational cost. To validate our approach, we first train a high-capacity teacher model on the BLAZE [39] dataset, which contains imagery from natural disasters such as wildfires. Then, we train a compact student model using a distillation strategy that preserves the teacher's retrieval performance. Our results show that the student model achieves comparable accuracy while being more suitable for real-time deployment.

AUTH has evaluated the proposed class-based retrieval method on the images from the Blaze wildfire dataset. The BLAZE dataset captures a diverse range of Natural Disaster Management (NDM) scenarios, featuring imagery related to wildfires. To address the limitations of conventional retrieval pipelines in NDM, AUTH introduced a novel approach based on representation distillation. Specifically, a Structurally-Aware Distillation Loss (SADL), which combines metric learning, feature correlation modeling, and knowledge distillation, is proposed. SADL encourages a lighter student DNN to replicate both the embedding structural relationships and the semantic structure encoded in the feature space of a deeper, more expressive teacher DNN.

This method jointly optimizes three objectives: the Triplet Loss that ensures that the student embeds semantically similar images closer together while pushing dissimilar ones apart, a Knowledge Distillation loss that transfers information from the teacher, encouraging the student to approximate the teacher's decision boundary and the global Correlation Loss that aligns the feature correlation maps between student and teacher, allowing the student to internalize global structural cues present in the teacher's intermediate feature space.

Unlike standard distillation, which focuses on aligning output logits, our approach directly optimizes the geometry of the student's latent space, making it more suitable for retrieval scenarios where relational understanding among samples is significantly important. To evaluate the efficacy

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

of our method, we trained a ResNet50 student model under the guidance of a ResNet101 teacher, using a curated triplet mining strategy with hard negatives and correlation-based regularization. An example of retrieved results from the Blaze Dataset is depicted in Figure 1. In Table 5, we present the retrieval performance on the original Blaze dataset. Additionally, to enhance the robustness of our method, we enriched the Blaze dataset with randomly sampled images, expanding it significantly to a total of 98,055 images; the corresponding retrieval results on this expanded dataset are provided in Table 6.

Figure 15. Result of the pipeline in the Blaze Dataset. The first column displays query images depicting various wildfire scenarios, including night-time and aerial views. The subsequent columns show the top three retrieved neighbor images based on visual similar

Table 12. Comparison between the teacher model (ResNet 101) and the student model (ResNet 50) trained under the examined scenarios on the Blaze Dataset.

Method	Triplet Accuracy	Retrieval mAP	Classification Accuracy
Teacher:	76.01%	0.7198	80.28%
		Student	
KD loss	72.79%	0.7079	81.58 %

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

KD and Triplet loss	75.75%	0.7280	81.49%
SADL	**86.12%**	**0.7351**	**83.71%**

Table 13. Comparison between the teacher model (ResNet 101) and the student model (ResNet 50) trained under the examined scenarios on the enriched Blaze Dataset.

Method	Triplet Accuracy	Retrieval mAP	Classification Accuracy
Teacher:	92.01%	0.788	75.28%
Student			
KD loss	81.33%	0.7721	74.54 %
KD and Triplet loss	93.02%	0.7919	75.38 %
SADL	93.05%	0.808	75.23 %

The results show that the AUTH structural distillation strategy improves the retrieval quality (measured via mAP) while reducing the computational footprint of the model, thus enabling deployment in resource-constrained or real-time disaster response systems.

5.7 Deployment and Scalability Considerations

The Geovisual Analytics module is engineered for robust, scalable, and maintainable deployment within the TEMA platform, leveraging modern cloud-native practices including containerization and orchestration.

Containerization and Registry Management:

The foundation of the deployment strategy is containerization. Each distinct layer or service within the Geovisual Analytics module (e.g., gateway, ETL workers, persistence helpers) is packaged as a lightweight Docker image. This approach ensures that each image carries only the essential subset of dependencies required for its specific function, promoting isolation, consistency across

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

environments, and efficient resource utilization. These Docker images adhere to semantic versioning conventions (e.g., <major>.<minor>.<patch>), ensuring traceability and reproducibility for each release.

The deployment pipeline automates the build, tagging, and publication of these Docker images to a private container registry, specifically GitHub Container Registry (GHCR) under the `he-tema/tema-geovisual-analytics` namespace. This automated process is typically triggered by version increments tracked in the environment configuration, streamlining the release cycle.

Orchestration with Kubernetes and Helm:

The containerized services are deployed and managed within a Kubernetes cluster, which for the TEMA project is orchestrated by the University of Messina (UNIME) using Rancher. Kubernetes provides the robust orchestration capabilities necessary for managing the lifecycle of these containerized applications.

Deployment manifests, which define how the application components run on Kubernetes, are managed using Helm charts. Helm allows for templating Kubernetes resources, making it straightforward to configure, deploy, upgrade, and rollback the Geovisual Analytics services seamlessly. Rancher then orchestrates the rollout of new container image versions within the Kubernetes cluster, typically employing controlled rolling updates to ensure zero downtime during deployments or updates.

Configuration parameters critical for the component's operation, including environment variables, sensitive information like API keys (secrets), and resource limits (CPU/memory), are managed externally using Kubernetes ConfigMaps and Secrets. This practice decouples configuration from the application code, enhancing security and flexibility.

Scalability Strategies:

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

The architecture is designed with scalability as a core tenet to handle varying workloads characteristic of emergency management scenarios, where data volume and processing demands can fluctuate significantly.

- **Horizontal Scalability:** The primary strategy is horizontal scaling. The analytical workloads, particularly within the Transformation phase utilizing "Analytics Building Blocks" and powered by Dask, are designed to be distributed across multiple worker pods in the Kubernetes cluster. As processing demand increases, Kubernetes can scale out the number of worker pods.
- **Autoscaling:** Kubernetes autoscaling policies are implemented, driven by metrics such as processing queue length and CPU load. This allows the cluster to automatically adjust the number of active pods for different services (especially the analytical workers) up or down, optimizing resource usage while maintaining performance.
- **Microservice Architecture:** The packaging of components into independent Docker images reflects a microservice-oriented approach. This allows different parts of the Geovisual Analytics module (e.g., data ingestion, specific analytical functions, data loading) to be scaled independently if needed, though the primary scaling focus is on the Dask-powered analytical workers.
- **Efficient Data Processing with Dask:** The use of Dask for parallel processing of large geospatial datasets (both raster and vector) is fundamental to the module's ability to scale its computational capacity. Dask's ability to manage distributed computations across worker pods ensures that complex analyses can be performed efficiently even on very large data.

Performance Optimization Considerations:

While specific low-level code optimizations are ongoing, the architectural choices inherently support performance:

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

- **Parallel Processing:** Dask's parallel execution model is key to minimizing processing times for large datasets.
- **Optimized Building Blocks:** The "Analytics Building Blocks" are designed for efficient execution of common geospatial operations.
- **Near Real-Time Goal:** The entire deployment and processing workflow is geared towards achieving (near) real-time performance to provide timely insights for emergency response.

Reliability, Maintainability, and Observability:

Continuous integration (CI) workflows are in place, incorporating code quality checks, unit and integration tests, and security scanning of container images before they are approved for deployment. This enhances the reliability and security of the services running in the production environment. For observability and troubleshooting, log aggregation from the multiple Kubernetes pods (e.g., facilitated by Kubernetes-native logging mechanisms or tools like Kubetail for consolidated viewing) provides a centralized stream for easier monitoring of system behavior and rapid identification of anomalies.

6. Conclusions

Since its inception, the Geovisual Analytics component (Task T5.2) has successfully established a robust, scalable, and modular platform for advanced geospatial data processing and preparation for visualization within the TEMA project. Key achievements include the implementation of foundational building blocks crucial for Natural Disaster Management, such as OpenStreetMap searches for Points of Interest, efficient zonal statistics computations, GeoTIFF to Cloud Optimized GeoTIFF conversions for optimized data handling, and the critical distance_from_fire metric. These functionalities have been rigorously validated through extensive testing and integrated seamlessly with the TEMA Digital Enabler ecosystem, particularly the FIWARE Context Broker, enabling real-time data ingestion and event-driven processing essential for emergency response.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

The component has demonstrated high performance and reliability in handling large and diverse geospatial datasets. This is largely attributed to the successful leveraging of parallel and distributed computing frameworks, notably the Python libraries rasterio, rioxarray, and especially dask for task parallelization and workload balancing. Integration with Kubernetes and Rancher for orchestration has ensured smooth deployment and inherent scalability, while continuous integration pipelines maintain code quality, security, and reproducibility. The close collaboration with visualization components like the SmartDesk (T5.3) has already provided end-users with actionable insights derived from T5.2 outputs, significantly enhancing situational awareness and decision-making capabilities in simulated emergency scenarios, particularly for fire risk assessment.

Moreover, the modular building block architecture of the Geovisual Analytics module has proven highly effective. This design allows for the rapid development and deployment of new analytical capabilities without disrupting existing workflows, ensuring the component can adapt to evolving project needs and new data sources. The system's comprehensive provenance tracking and metadata management, facilitated by adherence to NGSI-LD standards, ensure traceability and reproducibility of all analytical outputs. Overall, these achievements firmly position the Geovisual Analytics component as a cornerstone of the TEMA platform's geospatial intelligence infrastructure, successfully bridging the gap between raw data, sophisticated analytics (from WP3/WP4), Digital Twin models (T5.1), and effective end-user visualization (T5.3).

Future development, as outlined in the roadmap, will focus on extending the analytical capabilities of the Geovisual Analytics component by incorporating human-centric data layers. Planned building blocks will integrate demographic information, including population counts, age distributions, and gender profiles within defined geographic areas. This extension aims to enrich risk assessment models by correlating hazard proximity metrics (such as distance from fire or flood zones) with population density and vulnerability factors. Specifically, upcoming releases will implement composite risk indices that combine spatial proximity with demographic risk factors to better prioritize emergency response efforts. These indices will support differentiated risk classifications based on age and gender, enabling targeted interventions for vulnerable groups such as children, the elderly, and pregnant women. The component will also enhance semantic data annotation to incorporate these new data domains seamlessly. In parallel, improvements to the processing engine will continue, with optimizations for real-time execution and further automation of data ingestion workflows. Usability enhancements in the SmartDesk interface will allow end-users to visualize and interact with these enriched risk metrics dynamically. Additionally,

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

feedback loops from end-users and emergency responders will guide iterative refinements, ensuring that upcoming releases remain closely aligned with operational needs. By expanding the scope of geovisual analytics to include human factors, the TEMA platform will provide a more holistic understanding of risk, ultimately improving emergency preparedness, response, and resilience.

References

[1] Andrienko, G., Andrienko, N., Jankowski, P., Keim, D., Kraak, M. J., MacEachren, A., & Wrobel, S. (2007). Geovisual analytics for spatial decision support: Setting the research agenda. *International journal of geographical information science*, *21*(8), 839-857.

[2] Dykes, J., MacEachren, A. M., & Kraak, M. J. (2005). Exploring geovisualization. Elsevier.

[3] Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., ... & Cheng, T. (2016). Geospatial big data: A review of concepts, methods, and applications. ISPRS Journal of Photogrammetry and Remote Sensing, 115, 70-87.

[4] Singleton, A., & Arribas-Bel, D. (2021). Geographic data science. Geographical Analysis, 53(1), 61-75.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

[5] Kinkeldey, C., MacEachren, A. M., & Schiewe, J. (2017). How to assess the communicative fit of visualisations of geospatial uncertainty for supporting decision-making? A systematic review. Cartography and Geographic Information Science, 44(5), 370-386.

[6] Padmanabhan, A., Wang, S., Sha, D., & Liu, Y. (2019). A review of provenance in a Big Data-driven cyberGIS environment. ISPRS International Journal of Geo-Information, 8(2), 72.

[7] Janowicz, K., Krisnadhi, A. A., Hu, Y., Hitzler, P., Li, W., Vardeman, C., ... & Kolas, D. (2019). SOSA/SSN: The W3C OGC spatial data on the web joint working group. Semantic Web, 10(1), 9-19.

[8] Mai, G. (2016). Challenges in geospatial knowledge representation. In Workshop on AI for the geo-spatial world (GeoAI'16) (pp. 1-4).

[9] Kounadi, O., Ristea, A., Leitner, M., & Langford, C. (2020). A systematic review on the use of publicly available information for privacy-breaching inference in the context of location-based services. Transactions in GIS, 24(2), 257-285.

[10] Xu, T., He, L., Li, S., & Hu, M. (2017). Location privacy-preserving for big data in location-based services: A comprehensive survey. IEEE Access, 5, 1574-1595.

[11] Liu, Z., Cui, Y., Zhang, K., & Wang, J. (2021). A survey on visual analytics of spatio-temporal data. Visual Informatics, 5(1), 1-20.

[12] Demšar, U., Reades, J., Manley, E., & Batty, M. (2015). Using geovisual analytics to interactively explore real-time public transport performance data. Cartographica: The International Journal for Geographic Information and Geovisualization, 50(2), 116-130.

[13] Fox, P., Hendler, J., & Lee, T. (2020). Enabling FAIR Data in Earth and Space Sciences. EOS, Transactions American Geophysical Union, 101.

[14] Sacha, D., Senaratne, H., Kwon, B. C., Ellis, G., & Keim, D. A. (2019). The role of uncertainty, awareness, and trust in visual analytics. IEEE Transactions on Visualization and Computer Graphics, 25(1), 979-989.

[15] Keim, D., Kohlhammer, J., Ellis, G., & Mansmann, F. (Eds.). (2020). Solving problems with visual analytics. Eurographics Association.

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

[16] Gomes, V. C. F., Vinhoza, T. T. V., Paiva, A. C. D., & Albuquerque, J. P. D. (2018). A systematic review of the use of microservices architecture in GIS. ISPRS International Journal of Geo-Information, 7(12), 467.

[17] QGIS: https://qgis.org

[18] GeoNode: https://geonode.org

[19] NASA Worldview: https://worldview.earthdata.nasa.gov

[20] Kepler.gl: https://kepler.gl

[21] Google Earth Engine: https://earthengine.google.com

[22] PostGIS: https://postgis.net

[23] GDAL: https://gdal.org

[24] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. CVPR.

[25] Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. ICML.

[26] Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. MICCAI.

[27] CesiumJS: https://cesium.com/cesiumjs

[28] OpenLayers: https://openlayers.org

[29] AWS SageMaker Geospatial: ttps://docs.aws.amazon.com/sagemaker/latest/dg/geospatial.html

[30] Azure IoT Edge: https://azure.microsoft.com/en-us/services/iot-edge/

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

[31] AWS Greengrass: https://aws.amazon.com/greengrass/

[32] Khan, S. M. H., Hussain, A., & Alshaikhli, I. F. T. (2012, November). *Comparative study on content-based image retrieval (CBIR)*. In *2012 International Conference on Advanced Computer Science Applications and Technologies (ACSAT)* (pp. 61-66). IEEE.

[33] Arandjelovic, R., Gronat, P., Torii, A., Pajdla, T., & Sivic, J. (2016). *NetVLAD: CNN architecture for weakly supervised place recognition*. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5297-5307).

[34] Ali-Bey, A., Chaib-Draa, B., & Giguere, P. (2023). *MixVPR: Feature mixing for visual place recognition*. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision* (pp. 2998-3007).

[35] Berton, G., Masone, C., & Caputo, B. (2022). *Rethinking visual geo-localization for large-scale applications*. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 4878-4888).

[36] Berton, G., Trivigno, G., Caputo, B., & Masone, C. (2023). *EigenPlaces: Training viewpoint-robust models for visual place recognition*. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 11080-11090).

[37] Hausler, S., Garg, S., Xu, M., Milford, M., & Fischer, T. (2021). *Patch-NetVLAD: Multi-scale fusion of locally-global descriptors for place recognition*. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 14141-14152).

[38] Keetha, N., Mishra, A., Karhade, J., Jatavallabhula, K. M., Scherer, S., Krishna, M., & Garg, S. (2023). *AnyLoc: Towards universal visual place recognition. IEEE Robotics and Automation Letters, 9*(2), 1286-1293.

[39] Siavrakas, M., & Kitsos, F and I. Pitas (2024). *Blaze Fire Classification - Segmentation Dataset.* Zenodo. https://doi.org/10.5281/zenodo.11501836

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003

[40] Barz, B., Schröter, K., Münch, M., Yang, B., Unger, A., Dransch, D., & Denzler, J. (2019). *Enhancing flood impact analysis using interactive retrieval of social media images.* arXiv preprint arXiv:1908.03361.

[41] Cisneros, I., Yin, P., Zhang, J., Choset, H., & Scherer, S. (2022). *ALTO: A large-scale dataset for UAV visual place recognition and localization.* arXiv preprint arXiv:2207.12317.

[42] Xu, W., Yao, Y., Cao, J., Wei, Z., Liu, C., Wang, J., & Peng, M. (2024). *UAV-VisLoc: A large-scale dataset for UAV visual localization.* arXiv preprint arXiv:2405.11936.

[43] Shen, T., Luo, Z., Zhou, L., Zhang, R., Zhu, S., Fang, T., & Quan, L. (2018, December). *Matchable image retrieval by learning from surface reconstruction.* In *Asian Conference on Computer Vision* (pp. 415-431).

[44] European Telecommunications Standards Institute (ETSI). (2024). *ETSI GS CIM 009 V1.8.1: NGSI-LD API.* https://www.etsi.org/deliver/etsi_gs/CIM/001_099/009/01.08.01_60/gs_cim009v010801p.pdf

This project has received funding from the European Union's HE research and innovation programme under the grant agreement No. 101093003