# D2.3 – Knowledge graph model and custom ontology report

| | |
|---|---|
| Deliverable ID | D2.3 |
| Deliverable Title | Knowledge graph model and custom ontology report |
| Work Package | WP2 |
| | |
| Dissemination Level | Public |
| | |
| Version | 1.0 |
| Date | 31.09.2025 (M12) |
| Status | Submitted |
| | |
| Lead Editor | RISC |
| Main Contributors | NOA, LINKS, MITIGA, EDGE |

Project funded in the Horizon Europe Programme HORIZON-EUSPA-2023-SPACE under Grant Agreement 101180172

## Document History

| Version | Date | Author(s) | Description |
|---|---|---|---|
| 0.1 | 2025-05-31 | RISC | Initial draft |
| 0.2 | 2025-08-29 | RISC | Draft for internal review to NOA and LINKS |
| 1.0 | 2025-09-31 | RISC | Submitted version |

## Table of Contents

# List of Figures



# List of Tables

# Acronyms

| Acronym | Explanation |
|---|---|
| API | Application Programming Interface |
| EC | European Commission |
| EO | Earth Observation |
| GIS | Geographical Information System |
| GraphRAG | Graph-based Retrieval-Augmented Generation |
| HTTP | Hypertext Transfer Protocol |
| KG | Knowledge Graph |
| NLP | Natural Language Processing |
| PDO | Protected Designations of Origin |
| PGI | Protected Geographical Indication |
| RDF | Resource Description Framework |
| REST | Representational State Transfer |
| SDK | Software Development Kit |
| SIS2B | Service d'Incendie et de Secours de la Haute-Corse |
| SPARQL | SPARQL Protocol and RDF Query Language |
| WCAG | Web Content Accessibility Guidelines |
| WP | Work Package |

# 1 Executive summary

This deliverable presents the initial design and conceptual framework for the UNICORN Knowledge Graph (KG) and its accompanying custom ontology, aimed at enhancing disaster risk and emergency management through integrated semantic technologies. It introduces the principles of knowledge graphs, graph models, and ontologies, while outlining their application to fuse diverse datasets—ranging from satellite earth observation imagery to socio-economic indicators—into a coherent, interoperable structure. The report also describes the strategy for ontology development and previews the planned user interfaces, including access through an Application Programming Interface (API), a graph exploration dashboard, and a chatbot, that will enable accessible and intelligent interaction with the UNICORN graph. This report establishes the foundation for transforming the UNICORN project's data ecosystem into a semantically enriched knowledge space that can support real-time insight generation and decision-making.

# 2 Scope of the document

This **Knowledge Graph Model and Custom Ontology Report** outlines the foundational concepts, methodologies, and planned developments for the UNICORN Knowledge Graph (KG), designed to enhance disaster risk and emergency management through integrated, semantic data representation. The report introduces key topics such as KGs, graph data models, and ontologies, setting the theoretical framework for the implementation of the UNICORN project's KG and associated services.

This report also provides an overview of the datasets identified for potential integration and the strategy for developing the UNICORN ontology, which will enable rich semantic relationships across diverse data sources. Furthermore, this report previews the intended user interfaces, including APIs, dashboards, and chatbot tools, that will facilitate effective interaction with the KG.

This report establishes the core scope and direction for the UNICORN KG development, serving as a guide for project stakeholders. The technical implementation of the KG is part of WP3 and specifically T3.4 - Knowledge graph implementation.

## 2.1 Applicable documents

Table 1: Applicable documents

| ID | Title | Version | Date |
|---|---|---|---|
| [AD. 1] | Grant Agreement incl. Description of Action | 1.0 | 23/09/2024 |
| [AD. 2] | Consortium Agreement | 1.0 | 01/10/2024 |

## 2.2 Reference documents

Table 2: Reference documents

| ID | Title | Version | Date |
|---|---|---|---|
| [RD.1] | D1.4 GDPR, Gender Balance and Ethical Issues Management | 1.0 | 12/03/2025 |
| [RD.2] | D1.5 Data Management Plan | 1.0 | 31/03/2025 |
| [RD.3] | D2.2 Service Architecture and Technical Specifications report | 1.0 | 31/05/2025 |

# 3 Introduction

Typically, the data, maps, and applications generated for natural hazard management are utilised in isolation, which is a common challenge in Earth Observation (EO) applications. The data is often used once for analytics and mapping, neglecting its potential for long-term knowledge creation and sharing. A Knowledge Graph (KG) provides a state-of-the-art solution for integrating and connecting multiple datasets, enabling users — even without specific domain expertise — to explore the data, ask a range of questions on the topic, and unlock greater value from the underlying information than is usually accessible through a simple database solution. In today's data-rich world, the number of questions that can be answered grows exponentially with the volume of available data. Traditional approaches often struggle to manage this complexity, especially when working with large, predefined data tables and intricate interrelationships across multiple datasets. KGs offer a powerful alternative by simplifying the querying process, even in complex, multi-source environments. They reduce query times and lower the technical barrier for users, making it easier to extract meaningful insights without needing deep expertise in query languages or database structure.

A KG was proposed to be developed within the UNICORN project to demonstrate the integration of diverse data sources such as satellite imagery, weather data, vegetation maps, and socioeconomic information, offering a holistic perspective on natural hazards and disaster risk. Built with links to the UNICORN services, the KG enhances interoperability across systems and stakeholders, helping to uncover complex relationships and dynamics from partner hazard information and open data. Functioning as a digital twin, the KG will enable decision-makers to derive actionable insights through real-time updates and scenario analysis. The UNICORN KG will also serve as a central repository for hazard-related knowledge, supporting collaboration among researchers, policymakers, and emergency response teams. Selected datasets will be ingested into the UNICORN KG to fuse and enrich existing knowledge. By embedding semantic relationships between entities, the KG goes beyond the capabilities of traditional databases, enabling comprehensive, context-aware queries across domains.

# 4 Knowledge Model Definition

## 4.1 Knowledge Graphs

KGs differ fundamentally from traditional databases in both structure and modelling. Instead of relying on rigid schemas, KGs employ a flexible, graph-based structure where data is represented as nodes (entities) and edges (semantic relationships between entities). The basic building block of a KG is a triple composed of a subject, a predicate, and an object, for example: *Sentinel-2 satellite - captures - multispectral images of the Earth's surface*. This structure supports the integration of complex, unstructured, and interconnected data, enabling more intuitive understanding of relationships and advanced semantic querying, including through Natural Language Processing (NLP).

KGs are particularly suited for managing complex, relationship-heavy datasets. They allow for efficient deep and wide traversals, which are often performance bottlenecks in relational databases due to their reliance on computationally-expensive joint operations. In contrast, the graph structure enables explicit modelling of interdependencies, making KGs inherently more scalable and adaptable for dynamic, heterogeneous data environments. KGs are widely used across domains for recommenders' systems, question answering, and information retrieval. For example, Bloomberg uses KGs in its infrastructure to deliver business and financial news; Amazon and Netflix enhance product and content recommendations using KGs; and Google employs a KG to enrich search results by linking entities such as people, places, and events.

Moreover, the structured and semantically rich information in KGs can significantly enhance the performance of generative AI-based question answering systems. By grounding outputs in factual, interconnected knowledge, KGs help reduce hallucinations and improve reasoning. One emerging method, GraphRAG (Graph-based Retrieval-Augmented Generation) (Edge et al. 2024), integrates KGs into the retrieval process to guide generative models more precisely, ensuring that responses are based not just on loosely retrieved documents, but on curated, semantically organised knowledge.

In disaster risk and emergency management, a KG offers a transformative step beyond traditional data analysis tools, enabling a more integrated and dynamic understanding of risk, exposure, and response. By connecting diverse datasets — from environmental sensors to infrastructure records and socio-economic indicators, a KG facilitates a holistic view of complex hazard scenarios. Possible applications include:

- Linking wildfire activity data with critical infrastructure and population density to assess immediate impact zones and prioritise emergency response.
- Combining flood forecasts with real-time traffic and road network data to optimise evacuation planning and route guidance.
- Identifying vulnerable assets and populations by intersecting hazard zones with health, demographic, and economic data to inform preparedness and targeted interventions.
- Connecting satellite observations with insurance records and damage assessments to accelerate post-disaster recovery planning and claims processing.

These kinds of insights made possible by the semantic structure of a KG, go well beyond the capabilities of traditional, siloed database systems. While similar analyses could, in principle, be performed using conventional GIS or relational databases, they typically require manual data integration and lack the semantic context needed to fully capture the complexity of disaster risk scenarios.

By explicitly modelling entities and their interrelationships, a KG enables dynamic, interoperable, and context-aware querying across diverse and evolving datasets. This facilitates deeper reasoning, faster integration of new information, and more adaptive, coordinated decision-making — all of which are critical in time-sensitive disaster risk and emergency management situations.

## 4.2 Graph Models

Different graph models are used to represent data in KGs, with the two most common being Resource Description Framework (RDF) graphs[1] and Property graphs. In a directed edge-labelled RDF graph (Figure 1), each edge represents a directional relationship between two nodes, with labels that define the nature of the relationship. This model is particularly effective for representing binary relationships such as "fire detected by satellite" or "forest type fuels fire", where the interaction between entities is clearly defined.

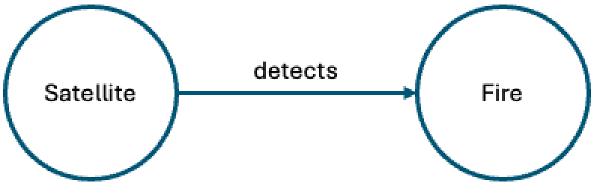


Figure 1: Example of an RDF graph model

The property graph model extends this structure by allowing both nodes and edges to carry additional attributes, or properties (Figure 2). These may include details like geographic coordinates, timestamps, fire intensity, or temperature, enriching the semantic context of both entities and their relationships.

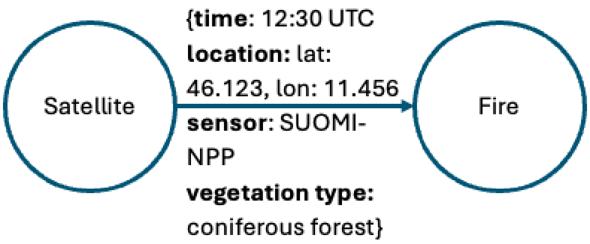


Figure 2: Example of a property graph model

The choice between graph models, as well as associated querying languages like Cypher (for property graphs) and SPARQL (for RDF graphs), depends on the specific requirements of the use case. During the service development, the graph model that best meets the requirements of UNICORN will be selected based on the specific characteristics of each use case and technology fit.

## 4.3 Ontologies

While a data model is a more specific, practical representation tailored to the requirements of a particular application or system, ontologies provide the high-level conceptual framework that describes domain knowledge in a structured manner. They capture domain knowledge at an abstract level, making it highly reusable and interoperable across various systems and applications. An ontology provides a formal structure of entities and relationships present in the graph and the rules governing their interactions. This ensures that data in a KG is meaningfully organized, consistently interpreted, and ready for complex querying and reasoning.

[1] RDF - Resource Description Framework, https://www.w3.org/TR/rdf11-concepts/

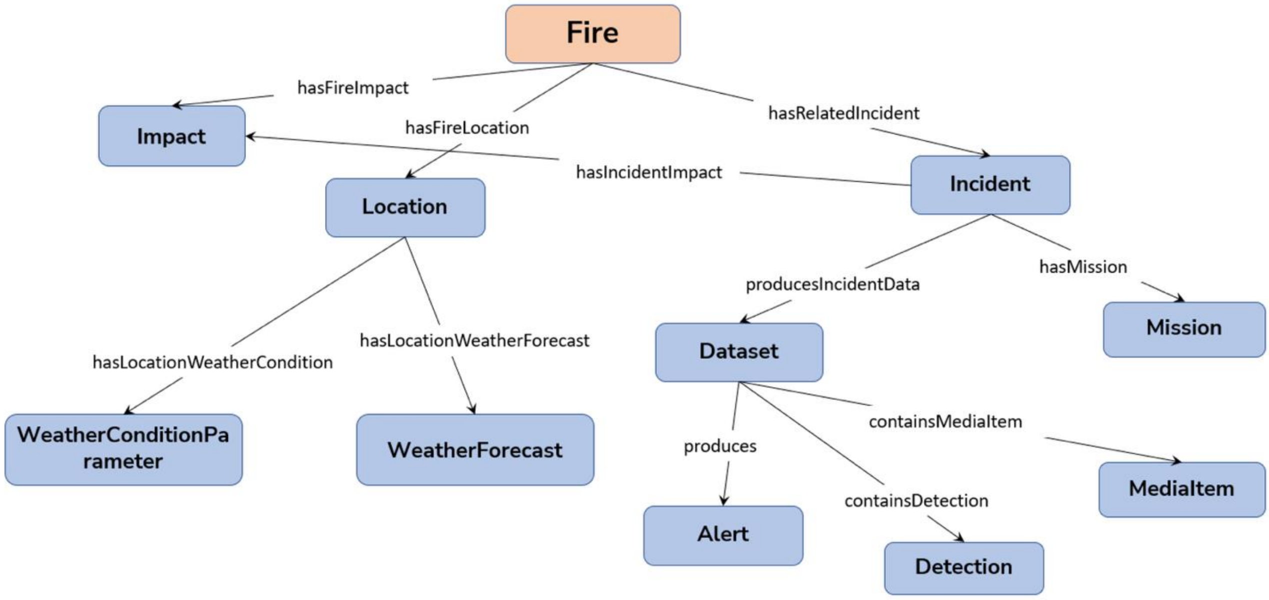


Figure 3: An example ontology presenting the main classes in the initial Onto-SAFE ontology[2]

In the context of disaster management, ontologies could define core entities such as hazard events, locations, weather conditions, or socio-economic conditions along with the relationships between them (Figure 3). By formalizing these entities and relationships, ontologies facilitate consistent data integration, enable complex querying and reasoning over the KG, and support a shared understanding of the domain across different systems and users. Building an ontology from scratch is time-consuming and complex. Thus, existing ontologies were examined to identify relevant concepts, relationships, and best practices that could be leveraged to accelerate the development process and ensure interoperability with other systems and datasets. The following pre-existing ontologies were investigated:

- SIADEX: Framework for forest fire combat strategies, using BACAREX knowledge base for planning (includes location, time, legal issues) (Asunción et al. 2005).
- EmergencyFire: Describes emergency fire response in buildings, with a taxonomy for tactical phases (Bitencourt et al. 2015).
- OntoFire: Geo-portal for wildfires, incorporating topography, vegetation, climate, and fire semantics (Kalabokidis et al. 2011).
- I-REACT: Disaster Risk Reduction (DRR) ontology bridging technology and end users, focusing on communication (I-REACT - Improving Resilience to Emergencies through Advanced Cyber Technologies, 2019).
- BeAWARE: Lightweight model for climate-related crisis management (https://github.com/beAWARE-project/ontology).
- NOA: Real-time fire monitoring using EO and sensor data, linked geospatial data, and semantic web technologies (Kyzirakos et al. 2014).

[2] Masa, P.; Kintzios, S.; Vasileiou, Z.; Meditskos, G.; Vrochidis, S.; Kompatsiaris, I. A Semantic Framework for Decision Making in Forest Fire Emergencies. Appl. Sci. 2023, 13, 9065. https://doi.org/10.3390/app13169065

- Semantic Sensor Network Ontology: General-purpose ontology for observation, actuation, and sampling interactions (https://www.w3.org/TR/vocab-ssn/).
- SAFERS: Conceptual map for forest fire disaster management, defining processes and data mapping for incidents and impacts (Masa, P. et al. 2023).

Other references were the KnowWhere Graph (Shimizu et al. 2025) and Bio Portal Fire Ontology for additional context (https://bioportal.bioontology.org/ontologies/FIRE?p=summary).

These examples are used not only to build-on but also to help UNICORN partners understand the ways in which their hazard information can be linked together within a graph context which is not always evident to domain experts.

# 5 Understanding the Available UNICORN Datasets

This chapter presents four UNICORN natural hazard scenarios to showcase the synergies between them. These natural hazard scenarios are for demonstration purposes, and the detailed design and implementation will be further developed and agreed during implementation of the project services.

## 5.1 Use Case Scenarios Overview

The UNICORN project addresses four distinct natural hazard scenarios, each presenting unique challenges and requiring specific data ecosystems:

- **Floods:** Focusing on flash flood phenomena in river basins, which necessitate rapid forecasting and subsequent impact assessment of dense urban and industrial environments.
- **Wildfires (Tourism/Evacuation):** Addressing the early detection, modelling of fire propagation, and the resulting impact on critical infrastructure, communities, and the natural environment.
- **Wildfires (Insurance):** High-resolution fire danger forecast and risk estimation aiming at reducing risk for the insurance sector.
- **Lava flows:** Centred on forecasting effusion paths from volcanic events and managing the consequent risk to populated areas and essential infrastructure.

Despite their phenomenological differences, these use cases share a common requirement for timely, actionable intelligence to support the decision-making processes of stakeholders such as civil protection authorities, critical infrastructure operators and insurances.

## 5.2 Core Data Integration Strategy

The central strategy for the UNICORN KG is to create a value-added analytical layer across the project's hazard forecasting systems. These forecasting systems provide essential spatio-temporal predictions of a physical hazard. The KG is engineered to contextualize this information, translating raw forecast data into actionable intelligence for decision-makers. The KG applications, enabled by the ontology defined in Chapter 6, are designed to answer questions regarding the vulnerability of exposed assets, the potential for cascading failures across critical infrastructures, and the overall risk to societal and economic functions. This transforms the direct outputs of forecasting models into a structured, queryable understanding of potential consequences.

## 5.3 Use Case - Flood impact on critical industrial infrastructure (Mandra)

**Objective:** To develop a KG application that assesses the risk and potential cascading failures for critical industrial infrastructure in the Mandra region, Attica, Greece.

**KG-enabled insights & stakeholder value:** This application moves beyond identifying inundated areas to providing specific, actionable intelligence. Civil Protection authorities can pre-identify industrial sites with hazardous materials at risk of leakage. Facility operators and emergency planners can anticipate cascading failures, such as power outages affecting multiple factories, allowing for proactive mitigation and resource staging. Stakeholders can pose complex queries that are vital during a crisis but difficult to be answered with traditional GIS or database systems.

**Example queries:**

- "Show all metallurgical facilities within the 2-day flood forecast zone."
- "Identify which sections of the main access road to the dockyards will be inundated with more than 0.3 meters of water."
- "Given the flood extent, list all industrial facilities classified under the Seveso Directive that are at risk and retrieve their emergency contact information."
- "Is the waste disposal site predicted to be flooded, and if so, what is the proximity to the nearest river stream for potential contaminant runoff?"
- "If the 'Elefsina' power substation is within the flood zone, generate a list of all dependent factories that will lose power."
- "Identify single points of failure: show me all bridges whose incapacitation would isolate more than three critical industrial facilities."

**Datasets for the application:** The KG ingests and integrates the following datasets to enable this analysis. The Ontology Mapping indicates how each dataset is represented in the KG:

Table 3: Datasets for Flood impact on critical industrial infrastructure UC

| Data Category | Dataset Description | Ontology Mapping |
|---|---|---|
| Hazard forecast | 2-day flood forecast (extent, depth). | mapped to flood:FloodForecast |
| Infrastructure | Industrial facility locations, types. | mapped to flood:IndustrialFacility (with specific subclasses) |
| Utilities | Electrical grid, road network topology. | mapped to flood:UtilityInfrastructure (e.g. PowerSubstation) |
| Dependencies | Facility-to-utility dependencies. | mapped to core:isDependentOn |

## 5.4 Use case - Wildfire impact on tourism and evacuation planning (Corsica)

**Objective:** To develop a KG application that assesses the dynamic risk to tourism-related assets and populations in Northern Corsica.

**KG-enabled insights & stakeholder value:** The application provides emergency managers (e.g. Service d'Incendie et de Secours de la Haute-Corse (SIS2B)) with a dynamic operating picture that accounts for the fluctuating tourist population. It allows for the rapid identification of high-capacity tourist sites under threat and, crucially, an analysis of the viability of evacuation routes. This supports more effective alert dissemination, safer evacuation planning, and better protection of both residents and visitors.

The KG will enable end-users to answer critical questions concerning the safety of both residents and the large transient tourist population.

**Example queries:**

- "Enumerate all registered campsites and resorts that intersect with the 6-hour probabilistic fire spread perimeter."
- "Identify all hiking trails and designated tourist areas located within 2 kilometres of the active fire front."
- "Calculate the total tourist accommodation capacity (e.g., number of beds, campsite pitches) within the areas designated as high-risk by the nowcast."
- "List all tourist accommodations that are downwind from the current fire, considering the latest meteorological forecast data for wind speed and direction."
- "Given the predicted fire progression, identify all primary and secondary evacuation routes from the coastal tourist areas that are projected to be compromised or have their accessibility reduced."
- "Generate a list of high-density tourist locations (e.g., campsites with capacity > 500) and determine the nearest viable civil protection assembly points, accounting for predicted road closures."
- "Identify critical telecommunication towers whose potential damage from the fire would disrupt emergency alert broadcasts to specific tourist-heavy valleys or coastal zones."

**Datasets for the application:** This analysis is enabled by the integration of the UNICORN wildfire system's outputs with a curated set of tourism and infrastructure datasets.

Table 4: Datasets for Wildfire impact on tourism and evacuation planning UC

| Data Category | Dataset Description | Ontology Mapping |
|---|---|---|
| Hazard forecast (Input) | Fire spread nowcasts (perimeters) | mapped to wildfire:FireSpreadForecast |
| Infrastructure | Campsite, resort, hotel locations/capacity | mapped to wildfire:TourismInfrastructure, hasAccommodationCapacity |
| Population | Seasonal population estimates | mapped to core:GeographicEntity instances (e.g., municipalities), hasEstimatedPopulation |
| Transportation | Road network topology and classification | mapped to: core:RoadInfrastructure, wildfire:EvacuationRoute. |

## 5.5 Use Case - Parametric insurance and forestry risk portfolio management (Iberian Peninsula)

**Objective:** To develop a KG application to serve as a risk analysis and portfolio management tool for insurance providers offering parametric products for forestry assets.

**KG-enabled insights & stakeholder value:** This application provides insurance underwriters and risk analysts with a novel tool for data-driven pricing and real-time exposure monitoring. By integrating dynamic ignition probability with their specific portfolio, they can calculate portfolio-wide value-at-risk, identify clusters of correlated risk, and automate the monitoring of parametric triggers, leading to more accurate, efficient, and innovative insurance products. The KG is designed to query insurances' forestry assets in the context of dynamic wildfire risk.

**Example queries:**

- "From the active insurance portfolio, list all forestry parcels located in areas with a daily ignition probability forecast exceeding a threshold of 0.7."
- "Generate a daily watch list of policies and associated holders for assets in the top 5% of ignition risk zones."
- "For all exposed parcels, calculate a 'Probable Maximum Loss' metric by correlating the insured value of the timber with the forecast ignition probability and typical fire spread characteristics for the dominant fuel model."
- "Identify policies for which the risk profile has changed significantly due to recent changes in land use or proximity to new infrastructure (e.g., powerlines)."
- "Calculate the total insured value at risk for the entire Galicia portfolio based on today's fire danger forecast."
- "Identify geographic clusters of high-value, high-risk insured parcels that could be impacted by a single large-scale fire event, thus representing a significant correlated risk."

**Datasets for the application:** This functionality is predicated on the integration of the UNICORN system's fire danger product with the insurer's internal business data.

Table 5: Datasets for Parametric insurance and forestry risk portfolio management UC

| Data Category | Dataset Description | Ontology Mapping |
|---|---|---|
| Hazard forecast (Input) | Daily ignition probability maps | Output of the UNICORN Fire Danger Forecast System (Use Case 2B) |
| Portfolio data | Policy details, insured values, triggers | mapped to insurance:InsurancePolicy, insurance:PolicyHolder, insurance:ForestryParcel |
| Forest cadastre | Geospatial boundaries of forestry assets | mapped to insurance:ForestryParcel |

## 5.6 Use Case - Impact on high-value agriculture and protected designations of origin (Etna)

**Objective:** To develop a KG application that quantifies the direct and indirect economic impact of a lava flow on high-value agricultural assets, particularly those with Protected Designations of Origin (PDO) and Protected Geographical Indication (PGI).

**KG-enabled insights & stakeholder value:** The application provides regional planning authorities and agricultural consortia with a tool to assess the permanent economic and cultural loss from an eruption. It moves beyond simply mapping the lava flow to quantifying the specific impact on the "Etna DOC" brand, estimating the financial loss in production value, and identifying wineries and cooperatives whose supply chains are severed. This supports long-term economic planning and disaster relief efforts. The KG will enable a multi-layered analysis of agricultural impact, translating a physical lava flow forecast into a socio-economic consequence report. Alternative PDO and PGI for Etna region could be Monte Etna PDO (Olive Oil), Etna Prickly Pear (Ficodindia dell'Etna) PDO, or Etna Cherry PGI[3].

[3] https://www.parks.it/parco.etna/Eprodotti.php?

**Example queries:**

- "Calculate the total area of agricultural land inundated by the forecast lava flow."
- "List all vineyard parcels that intersect with the lava flow path and provide their registered size."
- "Determine the total area of land within the 'Etna Rosso DOC' protected zone that is forecast to be lost."
- "Calculate the percentage of the total 'Etna DOC' production area that would be permanently removed from use by this eruption scenario."
- "Estimate the total annual production value lost by combining the inundated vineyard area with regional average yield and grape price data."
- "Identify all wineries and cooperatives that source a significant portion of their grapes from the impacted parcels to anticipate supply chain disruptions."
- "List critical agricultural infrastructure (e.g., irrigation channels, primary access roads for harvesting) destroyed by the lava flow."

**Datasets for the application:** This advanced analysis requires the fusion of the UNICORN lava flow product with detailed agricultural, economic, and administrative datasets.

Table 6: Datasets for Impact on high-value agriculture and protected designations of origin UC

| Data Category | Dataset Description | Ontology Mapping |
|---|---|---|
| Hazard forecast (Input) | Lava flow path forecast | mapped to lava:LavaFlowForecast |
| Agriculture | Cadastral data of agricultural parcels | mapped to lava:AgriculturalParcel (e.g. vineyard) |
| Protected zones | Official boundaries of PDO/PGI zones | mapped to lava:ProtectedGeographicIndication, isLocatedWithin |
| Value chain | Location of wineries, cooperatives | mapped to lava:AgriFoodFacility |

## 5.7 Cross-Use Case Synergies

While the above use cases are domain-specific, the underlying data and analytical needs reveal relevant synergies. A core strength of the KG approach is the ability to reuse and connect these shared concepts:

- **Critical infrastructure**: Transportation and utility networks (road networks, power substations) are fundamental to assessing cascading failures in the flood use case and evacuation logistics in the wildfire and lava flow use cases.
- **Geospatial and administrative boundaries**: All use cases rely on a common base layer of administrative boundaries, elevation models, and land cover data for geospatial context and analysis.
- **Stakeholder data:** Information about Civil Protection authorities, municipalities, and infrastructure operators can be modelled once and reused across all applications.

## 5.8 Summary of Use Cases, Insights, and Stakeholders

The following table summarizes the value proposition of the UNICORN KG across the four use case applications.

Table 7: Value proposition of the UNICORN KG across the four use case applications

| Hazard type | Key datasets | KG-enabled insights | Primary stakeholders |
|---|---|---|---|
| Floods | Hazard forecast, industrial facilities, utility networks | Cascading failure analysis, hazardous material risk identification | Civil Protection, industrial facility managers, emergency planners |
| Wildfires (Tourism) | Hazard forecast, tourism infrastructure, population estimates, road networks | Dynamic risk to tourists, evacuation route viability analysis | Fire & rescue services, Civil Protection, tourism operators |
| Wildfires (Insurance) | Hazard forecast, insurance portfolios, forest cadastre | Portfolio-wide exposure monitoring, parametric trigger automation | Insurance underwriters, risk analysts |
| Lava Flows | Hazard forecast, agricultural cadastre, PDO zones, economic data | Economic loss calculation, impact on protected brands, supply chain disruption | Regional planning authorities, Agricultural consortia, Civil Protection |

# 6 The UNICORN Ontology

This chapter provides the formal specification of the UNICORN ontology, i.e. the semantic schema that enables the KG applications described in Chapter 5. It also defines the types of entities, their properties, and the relationships between them, creating a structured blueprint for integrating and reasoning over interconnected data.

## 6.1 A Modular Design Approach

The ontology provides the formal structure that enables the KG applications. A single, monolithic ontology would be unwieldy; therefore, the UNICORN ontology is structured in layers:

- A core ontology that defines high-level, generic concepts applicable across all use cases.
- Specific risk modules for each use case e.g. flood risk, wildfire risk, that extend the core ontology with concepts unique to that hazard.

This modular approach ensures that the ontology directly enables the applications described in Chapter 5 by allowing a given application to load only the relevant modules while ensuring all share a common semantic foundation.

## 6.2 Ontology Visualization and Architecture

To provide a high-level overview, the ontology's architecture can be visualized as a series of interconnected layers. Figure 4 illustrates this structure.

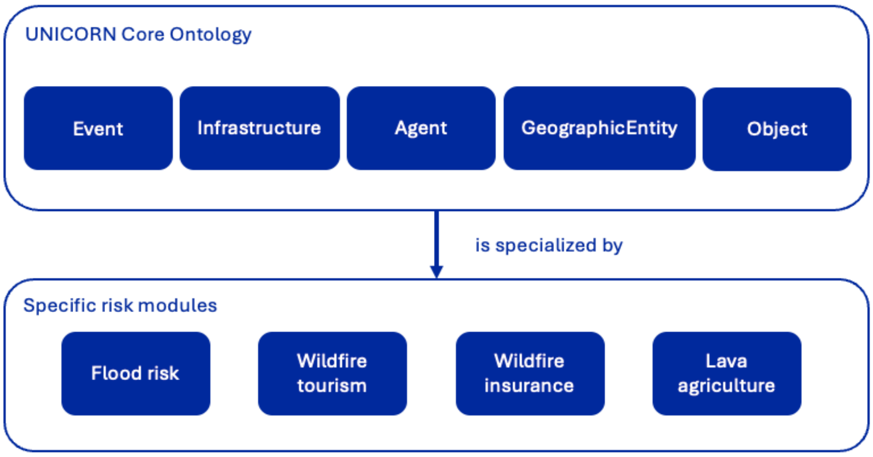


Figure 4: UNICORN core ontology and specific risk modules

This diagram shows the core ontology as the central component. It relies on external standards for foundational concepts. The specific risk modules extend the core ontology with domain-specific classes, inheriting its structure and ensuring overall consistency.

## 6.3 The UNICORN Core Ontology

### *6.3.1 The Flood Risk Ontology*

This module extends the core ontology to specifically model the domain of flood risk to industrial infrastructure in Mandra.

**Flood-specific classes:**

- Flood (subclass of core:Event)
- FloodForecast (subclass of core:InformationObject)
- IndustrialFacility (subclass of core:Infrastructure): The central class for this use case.
    - Subclasses: OilRefinery, Factory, WasteDisposalSite, Dockyard.
- UtilityInfrastructure (subclass of core:Infrastructure):
    - Subclasses: PowerSubstation, ElectricalLine, Road, Bridge.
- HazardousMaterial: A substance whose potential release poses a risk.
- ImpactStatement: A class to represent a derived conclusion from the graph (e.g., "Facility_A will lose power").

**Flood-specific properties:**

- stores (Object Property): Links an IndustrialFacility to a HazardousMaterial (e.g., a refinery stores crude oil).
- hasPredictedImpact (Object Property): Links a FloodForecast to an Infrastructure element, signifying it is within the predicted hazard zone.
- hasPredictedWaterDepth (Data Property): A numerical attribute of the hasPredictedImpact relationship, specifying the forecast water level at that location.
- crosses (Object Property): Links a Road or Bridge to a River.

### *6.3.2 The Wildfire Risk Ontology*

This module extends the core ontology with the specific classes and properties required to formally represent the domain of wildfire risk with a focus on tourism and public safety in Corsica.

**Wildfire-specific classes:**

- Wildfire (subclass of core:Event)
- FireSpreadForecast (subclass of core:InformationObject)
- TourismInfrastructure (subclass of core:Infrastructure)
    - Subclasses: Campsite, Resort, Hotel.
- RecreationalArea (subclass of core:GeographicEntity)
    - Subclasses: Beach, HikingTrail, NaturalPark.
- EvacuationRoute (a specialized subclass of Road, which is itself a subclass of core:Infrastructure)
- AssemblyPoint: A designated safe location for evacuation, subclass of core:GeographicEntity.
- PopulationGroup: A class to represent distinct population segments.
    - Subclasses: ResidentPopulation, TouristPopulation.

**Wildfire-specific properties:**

- hasAccommodationCapacity (Data Property): A data property of TourismInfrastructure to specify its size (e.g., number of persons).
- hasEstimatedPopulation (Data Property): A data property of a GeographicEntity (e.g., a municipality) to hold a value for a specific PopulationGroup.
- isDownwindFrom (Object Property): A topological and meteorological relationship between a GeographicEntity and a Wildfire.
- threatens (Object Property): A relation between a FireSpreadForecast and an Infrastructure or GeographicEntity.
- servesAsEvacuationRouteFor (Object Property): A relation linking an EvacuationRoute to a specific GeographicEntity or TourismInfrastructure it serves.

### *6.3.3 The Wildfire Insurance Risk Ontology*

This module extends the core ontology to formally model the concepts and relationships specific to the domain of wildfire insurance, risk assessment, and forestry management.

**Insurance-specific classes:**

- WildfireDangerForecast (subclass of core:InformationObject): Represents the daily ignition probability map produced by the UNICORN system.
- ForestryParcel (subclass of core:Infrastructure): Represents a managed land asset with economic value.
- InsurancePolicy (subclass of core:InformationObject): Represents the contractual agreement between the insurer and the policyholder.
    - Subclasses: ParametricInsurancePolicy, TraditionalIndemnityPolicy.
- PolicyHolder (subclass of core:Agent): The entity (person or organization) that owns the policy.
- ForestStand: A component of a ForestryParcel characterized by a specific tree species or age, representing a unit of value and risk.

**Insurance-specific properties:**

- covers (Object Property): A relation linking an InsurancePolicy to the ForestryParcel(s) it protects.
- hasInsuredValue (Data Property): A data property of an InsurancePolicy specifying its monetary value.
- hasParametricTrigger (Object Property): Links a ParametricInsurancePolicy to a specific condition (e.g., an FWI threshold) that would trigger a payout.
- isOwnedBy (Object Property): Links a ForestryParcel to a PolicyHolder.
- hasIgnitionProbability (Data Property): A data property of a ForestryParcel, updated daily from the WildfireDangerForecast, holding the numerical probability value.
- containsStand (Object Property): Links a ForestryParcel to one or more ForestStand entities.
- hasDominantSpecies (Object Property): Links a ForestStand to a tree species classification.

### *6.3.4 The Lava Flow and Agriculture Risk Ontology*

This module extends the core ontology to capture the specific concepts of agricultural land, value chains, and the unique, permanent nature of lava flow impacts.

**Lava flow & agriculture-specific classes:**

- LavaFlow (subclass of core:Event)
- LavaFlowForecast (subclass of core:InformationObject)
- AgriculturalParcel (subclass of core:Infrastructure): Represents a specific plot of land used for cultivation.
  - Subclasses: Vineyard, CitrusOrchard, OliveGrove.
- ProtectedGeographicIndication: A class to represent an official, geographically-defined quality scheme like a PDO or PGI.
- AgriFoodFacility (subclass of core:Infrastructure):
  - Subclasses: Winery, Cooperative.
- AgriculturalProduct: Represents the output of a parcel (e.g., grapes, wine).

**Lava flow & agriculture-specific properties:**

- isInundatedBy (Object Property): A relation that links an AgriculturalParcel or other infrastructure to a LavaFlowForecast, signifying permanent loss.
- isLocatedWithin (Object Property): A crucial geospatial relation linking an AgriculturalParcel to a ProtectedGeographicIndication zone.
- produces (Object Property): Links an AgriculturalParcel to the AgriculturalProduct it yields.
- isSuppliedBy (Object Property): A key dependency relationship linking an AgriFoodFacility (like a Winery) to the AgriculturalParcel(s) it sources from.
- hasAverageYield (Data Property): A data property of an AgriculturalParcel storing its typical production volume per hectare.
- hasAverageMarketValue (Data Property): A data property of an AgriculturalProduct storing its typical price per unit.

## 6.4 Ontology Instantiation and Reasoning Capabilities

The ontology is not merely a static model; it is an active component that enables the KG to ingest data and infer new knowledge.

**An example of ontology instantiation:** Consider the flood use case: the process of transforming raw data into knowledge within the KG proceeds as follows:

1. **Forecast ingestion:** A new 2-day flood forecast is generated. An application creates a new instance of the flood:FloodForecast class in the KG. This instance is populated with data properties, such as its validity time and a link to its geospatial data (the flood polygon).
2. **Geospatial analysis:** The application uses a GeoSPARQL or Cypher query (geo:sfIntersects) to find all existing instances of flood:IndustrialFacility whose stored location geometries intersect with the new flood polygon.
3. **Relationship creation:** For each industrial facility found, a new edge is created in the KG. This edge connects the FloodForecast node to the IndustrialFacility node with the label:Predicted_Impact. Additional details about the predicted impact, such as severity or expected damage, can be stored as properties on the edge.
4. **Enrichment:** The predicted water depth from the forecast data is added as a property of this new hasPredictedImpact relationship.

Through this process, raw data (a flood map) is transformed into a structured, queryable statement of knowledge: "FloodForecast_123 has a predicted impact on IndustrialFacility_ABC with a water depth of 0.8 meters."

**The Reasoning Capabilities:** The ontology's structure enables the KG to infer new information that is not explicitly stated in the source data. The primary mechanism is reasoning over relationships, particularly for cascading effects.

**Cascading Failure Inference**: The core:isDependentOn property is key. A user does not need to know which power station serves a factory. They can ask a simple query: "List all industrial facilities that will lose power if the 'Elefsina' substation is flooded." The KG answers this by:

1. Identifying that the substation isAffectedBy the flood.
2. Traversing the graph to find all IndustrialFacility instances that are isDependentOn the affected substation.
3. Returning this list of facilities as the result of the inference.

This ability to traverse and reason over modelled dependencies is a core capability that distinguishes the KG from traditional databases.

# 7 Knowledge Graph Interfaces

This section describes the various interfaces through which users could interact with the UNICORN KG. The goal of these interfaces is to support different levels of technical expertise and use case requirements, from programmatic access to intuitive visual exploration. The design of these interfaces will be informed by stakeholder feedback and usability testing.

To achieve this goal, the UNICORN KG should be accessible through three complementary interfaces: an **API** for programmatic integration, a **Dashboard** for intuitive visual exploration, and a **Chatbot** for natural language interaction. Together, these provide flexible access mechanisms for both technical and non-technical users, ensuring the KG can support a wide range of disaster risk management needs across the UNICORN stakeholders and beyond.

## 7.1 The UNICORN Knowledge Graph Application Programming Interface (API)

The Application Programming Interface (API) will provide programmatic access to the UNICORN KG, enabling integration with external systems, automated data retrieval, and the development of custom applications. The UNICORN KG will provide multiple programmatic access pathways to support both developers and integrators. The primary mode of access will be through language-specific client Software Development Kits (SDKs), while a RESTful (Representational State Transfer) API will also be available for cases where HTTP-based interactions are more appropriate.

### *7.1.1 Client SDKs - Primary Access*

Client SDKs will be offered in several widely used programming languages, including Python, JavaScript, Go, Java, Rust, and Swift, ensuring that the KG can be easily integrated into diverse application environments. These SDKs will:

- Provide direct access to the KG engine, enabling low-latency query execution and transactional guarantees.
- Allow users to create and manage schemas, ingest data, and run queries expressed in Cypher.
- Expose idiomatic APIs tailored to each language, making integration smooth for developers.
- Support both synchronous and asynchronous workflows, so they can be used in anything from lightweight scripts to large-scale services.

By embedding directly into the developer's environment, these SDKs offer the most efficient and feature-rich way to interact with the UNICORN KG.

### *7.1.2 RESTful API - Secondary Access*

As a complement to the SDKs, the UNICORN KG will also expose a **REST-style API.** This option is designed for environments where embedding an SDK may not be convenient, such as:

- Integrations across system boundaries,
- Low-code or no-code platforms,
- Rapid prototyping, or
- Access by lightweight clients that only need occasional interactions.

The REST API will provide endpoints for common operations, such as:

- GET /status — server health and availability checks.
- GET /schema — retrieve the current data model (node and relationship types, properties).

- POST /cypher — execute queries via HTTP, returning structured JSON responses.

The API will support configurations such as read-only vs. read–write modes, and optional optimizations for large-scale queries such as pagination and asynchronous query execution.

### *7.1.3 Authentication, Documentation, and Developer Support*

Both the SDKs and REST API will be protected by robust **authentication and authorization** mechanisms, such as API keys or OAuth2 tokens, with role-based access control to safeguard sensitive data. Details related to OAuth2 implementation have been provided in the Services Architecture and Technical Specifications Report (D2.2).

To ensure usability, the interfaces will be supported by:

- **Comprehensive documentation** for each SDK, with language-specific code examples.
- **REST documentation** based on OpenAPI/Swagger[4], allowing users to understand and integrate the REST API
- **Sample workflows** demonstrating typical tasks in disaster risk management, such as identifying at-risk populations or analysing hazard propagation.

## 7.2 UNICORN KG Dashboard

The dashboard will serve as the primary visual interface for exploring and interacting with the UNICORN KG. It is intended to provide intuitive access for users who may not be familiar with graph query languages, enabling them to discover insights through direct visual exploration and guided interactions.

### *7.2.1 Dashboard Core Capabilities*

The dashboard will offer interactive tools to **query, filter, and visualize** the KG:

- **Graph visualisation:** Nodes and relationships can be displayed in an interactive canvas, allowing users to zoom, pan, expand connections, and follow paths through the graph.
- **Schema exploration:** Users can inspect the structure of the KG (node types, relationship types, properties) to understand the underlying data model.
- **Faceted search and filtering:** Filters by time, location, entity type, or other attributes help narrow down large graphs to relevant subsets.
- **Result views:** Query results can be presented as tables, charts, maps, or graph diagrams, depending on the analysis needed.

These capabilities will make it possible to start from a simple search (e.g., "flood risk in Region X") and then progressively refine the results by exploring related entities (e.g., infrastructure, population, supply chains).

### *7.2.2 Advanced Features*

To support analytical and decision-making workflows, the dashboard will also include:

[4] https://swagger.io/specification/

- **Geospatial integration:** Overlay graph entities (such as hazards, critical infrastructure, or vulnerable populations) onto interactive maps, enabling spatial analysis.
- **Time-based exploration:** Visualize changes over time, such as how risks evolve or how hazards affect different areas across different scenarios.
- **Collaboration features:** Shared queries, annotations, and exports (e.g., CSV, PNG, or JSON) will support multi-user workflows across teams.

#### *7.2.3 Usability and Accessibility*

The dashboard will be designed with **usability and inclusivity** in mind:

- **Guided query builders** will help non-technical users formulate queries without needing to know Cypher.
- **Interactive tutorials** and **tooltips** will guide first-time users in navigating the interface.
- **Accessibility compliance (WCAG)** will ensure the interface can be used by individuals with diverse needs.
- **Responsive design** will allow access from desktops, tablets, and mobile devices, making it usable in the field during emergency operations.

#### *7.2.4 Integration with the API*

The dashboard will act as a **client of the UNICORN KG API**:

- All queries and visualizations will be backed by the same API endpoints available to developers.
- This ensures consistency across programmatic and visual access modes.
- Advanced users will be able to switch between graphical query builders and raw Cypher queries, using the dashboard as both a beginner-friendly tool and a power-user interface.

#### *7.2.5 Benefits*

The dashboard will democratize access to the UNICORN KG, making complex graph structures **transparent, explorable, and actionable** for a wide audience—from technical analysts to field responders and decision-makers.

### 7.3 Chatbot

The chatbot will provide a **conversational interface** to the UNICORN KG, enabling users to obtain insights by writing **natural language queries** instead of Cypher. This lowers the technical barrier and allows non-specialists—such as field responders, policy planners, or local stakeholders—to interact directly with the KG without needing training in query languages.

#### *7.3.1 Chatbot Core Capabilities*

The chatbot will translate natural language input into structured graph queries and return responses in clear, interpretable formats. Key features include:

- **Natural language querying:** Users can simply ask questions such as *"Which hospitals are at risk of flooding in Region X?"* and receive fact-based answers.

- **Query translation:** The system will automatically convert user input into underlying Cypher queries, execute them against the KG, and return structured results.
- **Multi-modal responses:** Answers could be presented not only as text but also as **tables, charts, or small graph visualizations**, depending on the context.
- **Contextual guidance:** If a query is ambiguous, the chatbot will prompt follow-up questions to clarify intent.

### *7.3.2 Advanced Features*

To maximize usability and trust, the chatbot will support:

- **Knowledge grounding:** Every response will be backed by references to the KG, ensuring that outputs remain factual and auditable.
- **Hybrid LLM–KG approach:** A large language model (LLM) will be used to interpret user intent, while the KG provides the authoritative data.
- **Explainability:** Where possible, the chatbot will show the Cypher query it generated, giving advanced users insight into how their natural language was interpreted.
- **Domain-specific tuning:** The chatbot will be optimized for disaster risk management scenarios, ensuring accurate handling of terms like hazards, exposure, and resilience.
- **Integration with collaboration tools:** The chatbot could be embedded into external platforms, enabling quick access during crisis situations. The UNICORN demonstration platform is such a use case.

### *7.3.3 Usability and Accessibility*

The chatbot will be designed for **ease of use in high-pressure environments**:

- Available through both web and mobile interfaces,
- Optimized for short, direct questions as well as conversational exploration,
- Equipped with simple commands such as *"show related data," "summarize,"* or *"compare regions."*

### *7.3.4 Example Usage*

The chatbot will support a variety of interaction styles that highlight its ability to replace technical queries with natural language questions:

- **Simple fact-finding:** A user asks a straightforward question such as *"Which regions are most exposed to hazards?"* and receives a concise, data-backed answer.
- **Exploratory analysis:** A user poses a broad question like *"Show me the relationships between hazards and critical infrastructure"* and is guided through visual or tabular outputs.
- **Comparisons:** A user requests *"Compare vulnerability across two regions"* and receives a structured summary that highlights similarities and differences.

- **Follow-up refinement:** After receiving an initial answer, the user can naturally continue the conversation with clarifying questions such as *"Focus only on health facilities"* or *"Show changes over the last five years."*

### *7.3.5 Benefits*

The chatbot will make the UNICORN KG **accessible to the widest audience**, including those without technical expertise. By bridging natural language and structured graph data, it ensures that decision-makers can access timely, accurate, and explainable insights during both planning and emergency response.

# 8 Conclusions

The **goal** of this report was to outline the basis for the development and implementation strategies of the UNICORN KG and related services. It is used as a KG primer to help partners understand the concepts and presents the different options available for implementing the UNICORN KG services. The report also highlighted the benefits of the KG to support interoperability, data integration, and decision-making in disaster risk and emergency management.

Knowledge graphs represent data as interconnected entities and relationships which differ in the manner they are stored and queried compared to traditional databases. It is important to realise that a KG is not simply transforming a traditional geospatial database but requires significant input from data owners, domain specialists and partner cooperation to identify the relevant relationships, access to the UNICORN data, and process it to be ingested into the KG infrastructure. While this KG related discussion has already begun, the implementation phase will require significantly more interaction and partner inputs to successfully develop the UNICORN KG and associated services:

- While a KG is expected to deliver superior handling of complex relationships and semantics, there is an increased **complexity** related to the setup and querying of the data. This potential learning can be significant and has been a barrier to adoption in many cases. Therefore, the implementation team will provide tools to help **increase adoption** across partner organisations to demonstrate the improvements over traditional databases.
- The expectation of adopting and implementing the UNICORN KG is **faster query performance** for relationship-heavy tasks which is wanted in the context of emergency responses to natural hazard events, there is a real challenge for the implementation of such services from a maintenance and consistency perspective because the data ingestion must be well managed to ensure data integrity. The goal is to demonstrate queries across scenarios with large-scale interconnections.
- In the case of the UNICORN KG, the goal is to provide **advanced analytics** for **improved decision-making** insights by incorporating semantics and ontologies to handle the dynamic hazard datasets. To deliver on this promise, there is a requirement for resources and expertise compared to traditional databases. This will be co-developed based on the UNICORN partner expert recommendations.

The current report is the basis for the implementation of the UNICORN KG and places all project partners on the same level of understanding to decide on the implementation details. The UNICORN KG and its custom ontology represent a transformative approach to disaster risk and emergency management by integrating diverse datasets such as EO imagery, hazard models, socio-economic indicators, and critical infrastructure into a semantically enriched, interoperable framework. The proposed interfaces, including an API, interactive dashboard, and natural language chatbot, further democratize access to the KG, enabling both technical and non-technical users to explore and extract value from the data seamlessly.

This deliverable has outlined the conceptual foundation, graph models, and potential ontology designs that allow the UNICORN KG to address the four critical use cases: flood impacts on industrial infrastructure, wildfire risks to tourism and evacuation planning, wildfire-related insurance portfolio management, and lava flow effects on high-value agriculture. This document will be used to develop the operational deployment of the UNICORN KG as part of T3.4 to finally support interoperability, data integration, and decision-making in disaster risk and emergency management.

# 9 References

Bitencourt, K., Durão, F., & Mendonça, M. (2015, October). Emergencyfire: An ontology for fire emergency situations. In *Proceedings of the 21st Brazilian Symposium on Multimedia and the Web* (pp. 73-76).

De la Asunción, M., Castillo, L., Fdez-Olivares, J., García-Pérez, Ó., González, A., & Palao, F. (2005). SIADEX: An interactive knowledge-based planner for decision support in forest fire fighting. *Ai Communications*, *18*(4), 257-268.

Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., ... & Larson, J. (2024). From local to global: A graph rag approach to query-focused summarization. *arXiv preprint arXiv:2404.16130*.

Kalabokidis, K., Athanasis, N., and Vaitis, M.: OntoFire: an ontology-based geo-portal for wildfires, Nat. Hazards Earth Syst. Sci., 11, 3157–3170, https://doi.org/10.5194/nhess-11-3157-2011, 2011.

Kyzirakos, K., Karpathiotakis, M., Garbis, G., Nikolaou, C., Bereta, K., Papoutsis, I., ... & Kontoes, C. (2014). Wildfire monitoring using satellite images, ontologies and linked geospatial data. *Journal of web semantics*, *24*, 18-26.

Masa P, Kintzios S, Vasileiou Z, Meditskos G, Vrochidis S, Kompatsiaris I. A Semantic Framework for Decision Making in Forest Fire Emergencies. *Applied Sciences*. 2023; 13(16):9065. https://doi.org/10.3390/app13169065

Shimizu, C., Stephen, S., Barua, A., Cai, L., Christou, A., Currier, K., ... & Zhu, R. (2024). The KnowWhereGraph ontology. *Journal of Web Semantics*, 100842.

End of document