The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume L-4/W1-2026
Free and Open Source Software for Geospatial (FOSS4G) 2026 – Academic Track "Bridging Geospatial Technology and Humanity",
30 August–5 September 2026, Hiroshima, Japan

Towards a Spatial Knowledge Mesh: A Metamodel for Federated and Interoperable Spatial Knowledge Graphs to Enable Geospatial Awareness, Integrity, Provenance, and Trust in Large Language Models

Nathan McEachen, Justin Smethie

TerraFrame, Inc., 75 Waneka Parkway Suite 210, Lafayette, Colorado, 80026 USA – (nmceachen, jsmethie)@terraframe.com

Keywords: Spatial Knowledge Mesh, Geo-GraphRAG, GeoAI, Common Geo-Registry, Integrity, Provenance and Trust.

Abstract

Large language models (LLMs) have the potential to make geospatial decision support more accessible to non-technical users and under-resourced settings in public health, disaster response, and infrastructure planning. Compared with traditional geospatial workflows requiring specialized expertise, manual data integration, preprocessing, and static mapping, LLM-based GeoAI systems can enable users to explore spatial questions through natural language. However, LLMs lack an internal representation of geographic features, spatial networks, and domain-specific semantic relationships, and therefore require structured geospatial knowledge at runtime to reason over interconnected spatial systems. Building on prior work that introduced the Spatial Knowledge Mesh (SKM) as an architecture for managing dependencies and interoperability between interlinked spatial knowledge graphs (McEachen, N., Lewis, J., 2023), this paper makes three contributions. First, it defines a metamodel for publishing interoperable spatial knowledge graphs (iSKGs) as governed, versioned, and temporally valid knowledge assets. Second, it describes how iSKGs can be federated through shared geographic abstractions, geo-ontologies, provenance metadata, and change propagation mechanisms. Third, it shows how this publication model enables Geo-GraphRAG workflows in which LLM-based systems can retrieve and reason over spatial relationships with integrity, provenance, and trust (IPT). By enabling governments, spatial data infrastructures, and other institutions to publish networks of geographic features as interoperable knowledge assets, the SKM can help lower-resourced settings create locally relevant GeoAI applications and make public geospatial data more effective for decision support.

1. Introduction

Large language models (LLMs) have created new possibilities for making geospatial decision support accessible to non-technical users. In domains such as disaster response, public health, transportation, and infrastructure planning, many important questions are spatially complex: Which communities may lose access to hospitals if a road network is flooded? Which schools are downstream of wildfire-prone watersheds? Which facilities depend on infrastructure located inside a hazard zone? Traditional geospatial workflows can answer such questions, but they often require specialized software, manual data integration, layer selection, preprocessing, and domain expertise. LLM-based GeoAI systems offer the possibility of allowing users to ask these questions through natural language.

However, LLMs do not maintain an intrinsic, up-to-date model of geographic features, spatial networks, or domain-specific relationships. They may synthesize text about roads, hospitals, schools, or hazards, but they do not natively know which specific features are connected, contained, dependent, adjacent, or exposed at a given place and time. This limitation is especially important for decision-support applications, where answers must be grounded in authoritative data, reproducible workflows, and traceable evidence (Figure 1). Recent work has also emphasized that LLMs and LLM-based agents face limitations when carrying out or verifying tasks beyond a certain level of complexity (Sikka and Sikka, 2025). For geospatial applications, this reinforces the need to reduce the reasoning burden placed on the language model by providing structured, queryable knowledge at runtime.

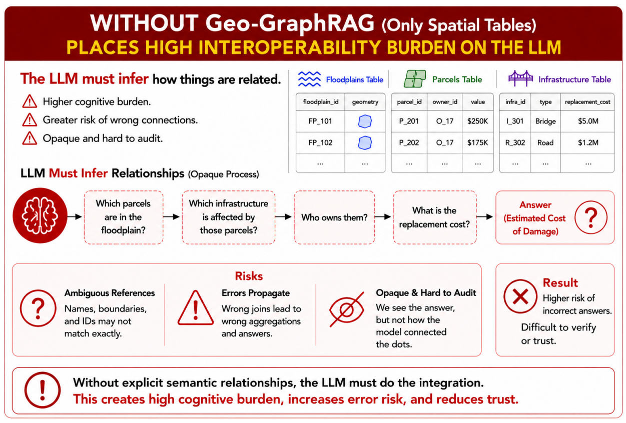

Figure 1. LLMs lack knowledge of feature networks

Knowledge graphs provide one mechanism for addressing this limitation. LLMs are often characterized as black-box models that can struggle to access and verify factual knowledge, while knowledge graphs explicitly represent entities and relationships in a machine-readable structure that can support inference and interpretability (Pan et al., 2024). Retrieval-augmented generation (RAG) extends LLMs by retrieving external information at runtime, but conventional vector-based RAG is best suited to local retrieval tasks and can struggle with questions requiring global understanding of relationships across a dataset. GraphRAG addresses this limitation by constructing graph-based indexes that support query-focused summarization and sensemaking over large text corpora (Edge et al., 2025). This paper extends that idea into the geospatial domain. Rather than relying on LLMs to infer spatial relationships from text, Geo-GraphRAG requires spatial knowledge graphs whose nodes represent geographic features and whose edges represent topological, semantic, temporal, or analytical relationships derived from geoinformatics data. These graphs allow LLM-based systems to translate natural-language questions into graph queries, retrieve structured spatial evidence, and generate responses grounded in explicit relationships rather than unsupported inference (Figure 2).

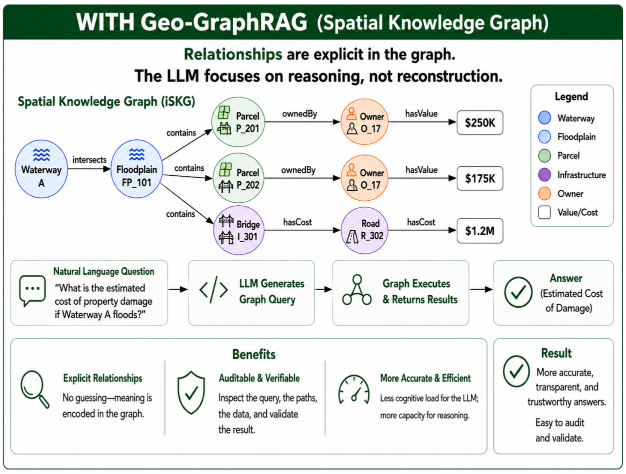

Figure 2. Geo-GraphRAG enables LLMs to focus on reasoning

This is especially important for addressing the complexity limitations of LLMs for geospatial decision support because many spatial questions are inherently compositional. Answering whether a community may lose access to hospitals during a flood, for example, may require combining road connectivity, hazard exposure, facility locations, administrative boundaries, population distributions, and temporal conditions. Recent work on LLM reasoning failures identifies compositional reasoning as a fundamental weakness: LLMs may be capable of handling individual components of a task while failing to integrate those components into a coherent inference, particularly as reasoning depth increases or irrelevant information is introduced (Song et al., 2026). For GeoAI applications, this suggests that complex tasks should not depend solely on implicit reasoning inside the language model. Instead, intermediate outputs can be externalized as structured, interoperable semantic relationships, or edges, between geo-objects—that is, geographic features represented as nodes in a graph. When one task produces a relationship such as a facility being exposed to a hazard, dependent on a road segment, or contained within an administrative area, that output can be persisted as a governed edge in a spatial knowledge graph (Figure 2). Subsequent tasks can then retrieve and reason over explicit relationships rather than reconstructing them from unstructured context, reducing the cognitive burden placed on the LLM and improving traceability across multi-step workflows.

Building on prior work that introduced the Spatial Knowledge Mesh as an architecture for automating interoperability and dependency management between spatial knowledge graphs published by version and period of validity (McEachen and Lewis, 2023), this paper advances the concept by defining a metamodel for interoperable Spatial Knowledge Graphs (iSKGs). The proposed metamodel supports governed publication, shared geographic abstractions, geo-ontologies, provenance metadata, temporal validity, versioning, and change propagation. Rather than relying on custom point-to-point integrations between systems, interoperability should occur at the point where spatial knowledge is published. This enables independently governed organizations to publish reusable knowledge assets that can be discovered, combined, and queried on demand.

The remainder of this paper describes how iSKGs can be maintained within an SKM and used to support Geo-GraphRAG workflows. The goal is to make networks of geographic features findable, interoperable, and usable by LLM-based applications, thereby helping governments, spatial data infrastructures, and other institutions make public geospatial data more effective for decision support and the public good.

2. Background

As discussed in our previous work (McEachen and Lewis, 2023), this approach originated in the disease intervention and elimination domain through efforts to develop a Common Geo-Registry (CGR). A CGR provides a governed source of truth for managing geographic features, their attributes, geometries, and relationships across organizations and information systems. This enables data from different sources to be contextualized in space and time, aggregated through geographic hierarchies, and linked through shared geographic objects. Temporal management is particularly important because geographic units may change, split, or merge over time (Ebener, 2022).

The prior work identified the need for a federated approach inspired by the Data Mesh architecture pattern, in which independently governed spatial knowledge graphs can be published, discovered, and integrated through shared geographic references and dependency management mechanisms. Representing CGR concepts as spatial knowledge graphs offers advantages in managing temporal evolution, semantic relationships, and dependencies among geographic features.

These concepts were subsequently applied to the development of Geo-GraphRAG infrastructure for LLM-based geospatial reasoning. Implementations with the U.S. Army Corps of Engineers (USACE) Civil Works Division and collaborative research through the Open Geospatial Consortium (OGC), supported by Natural Resources Canada and the United States Geological Survey, demonstrated the value of spatial knowledge graphs for providing structured geospatial context to LLM-enabled disaster response and resilience workflows (Open Geospatial Consortium, 2025).

However, these implementations were built on a platform architecture that was not designed as a fully federated Spatial Knowledge Mesh. While they demonstrated the value of Geo-GraphRAG over spatial knowledge graphs, they did not fully support independent publication, versioning, change propagation, and cross-organization integration of interoperable spatial knowledge graphs. This paper addresses that gap by presenting a new metamodel that incorporates lessons learned from prior CGR, SKM, and Geo-GraphRAG implementations to support a more complete realization of the Spatial Knowledge Mesh architecture pattern.

3. Semantic Standards for AI Readiness

Historically, semantic standards have been justified primarily by their ability to improve interoperability. In the age of AI, they offer an additional and potentially larger return on investment: they reduce the cognitive burden placed on both human analysts and AI systems. Every relationship that is explicitly modeled is a relationship that does not need to be rediscovered, inferred, or reconstructed at runtime. Rather than expending effort reconciling disparate representations of the same spatial reality, an LLM-based system can focus on reasoning over explicitly represented knowledge. This can support more accurate answers, greater transparency, improved auditability, and more effective use of computational resources. The value of semantic standards is therefore not limited to data exchange; it lies in creating reusable knowledge infrastructure that amplifies both human and artificial intelligence. In agentic workflows, semantic standards also make intermediate outputs more reusable and interoperable, allowing one task to produce structured knowledge that subsequent tasks can consume with less ambiguity.

3.1 Interoperability Up Front: Lessons from Terminologies and DGGS

Interoperability is most effective when designed into data at publication rather than reconstructed by each downstream consumer. Terminology systems demonstrate this principle by defining concepts, identifiers, relationships, and governance rules before data are exchanged, allowing independently developed systems to interpret meaning consistently.

Discrete Global Grid Systems (DGGS), standardized by the OGC, provide a parallel example for spatial data. By aligning aggregate and statistical datasets to standardized zone identifiers, DGGS reduces the burden of spatial integration and enables datasets from different domains to be combined through a common geographic reference framework.

Spatial Knowledge Graphs (SKGs) provide a complementary mechanism for representing networks of geographic features and relationships such as containment, adjacency, dependency, exposure, accessibility, and service coverage. However, most SKGs remain local or project-specific artifacts. No widely adopted equivalent to DGGS currently exists for publishing SKGs as interoperable, independently governed spatial knowledge assets.

This paper argues that interoperable Spatial Knowledge Graphs (iSKGs) should extend the principle of interoperability up front into feature-based spatial knowledge by publishing stable geo-object identities, explicit relationship semantics, provenance, temporal validity, versioning, and dependency metadata.

3.2 Why Local SKGs Are Not Enough

Public health provides a useful analogy for why spatial knowledge graphs require more than local implementation. Terminology services are critical in digital health because they enable independently developed systems to use clinical, public health, and administrative concepts consistently. By providing standardized code systems, value sets, and mappings, they reduce ambiguity and support data exchange, aggregation, interpretation, and reuse. This role is recognized in the emerging WHO Reference Architecture for Digital Public Infrastructure for Health, which includes a Terminology Service as a core digital public infrastructure for health.

Spatial knowledge graphs require similar semantic rigor, but differ from terminologies in an important way. Many terminologies are governed by global or national standards authorities, whereas spatial knowledge is often locally curated by the source authority responsible for a particular class of geographic features. A national mapping agency may govern administrative boundaries, a ministry of health may govern operational health areas, a utility may govern infrastructure networks, and local governments may govern roads, schools, or service zones. These authorities may also publish spatial knowledge that depends on features maintained by others.

A locally created SKG can support internal use, but it cannot function as shared interoperability infrastructure unless it is published with explicit semantics and governance metadata. Without common geo-object identities, defined relationship classes, temporal validity, provenance, versioning, and dependency information, independently published SKGs cannot be reliably interpreted or integrated. This becomes especially important for LLM-based GeoAI: an LLM may retrieve information from multiple graphs, but it cannot safely reason across them if meaning, authority, and temporal validity are implicit or inconsistent. To support cross-domain reasoning, SKGs must be published as interoperable spatial knowledge assets rather than isolated local graphs.

3.3 Fault Scenario: When Spatial Knowledge Is Not Interoperable

There are several fault scenarios that can arise when spatial knowledge is not published with sufficient semantic, temporal, and provenance structure. These include mismatches in feature identity, relationship semantics, source authority, ontology version, spatial resolution, and period of validity. This paper focuses on one illustrative scenario: a temporal mismatch between operational data and the geographic boundaries used to interpret that data.

Consider a district, D1, that reports 1,000 disease cases at time T0. At time T1, D1 is split into two districts: D1′ and D2. At time T2, the operational case data correctly records 300 cases in D1′ and 700 cases in D2. However, if an analytical system or LLM-based workflow continues to use the old boundary model, it may recognize only D1 and fail to represent D2. In that case, the 700 cases associated with D2 effectively fall outside the system's geographic frame of reference. A time-series comparison between T0 and T2 may therefore appear to show a decline from 1,000 cases to 300 cases, producing the false conclusion that disease burden has fallen by 70 percent (Figure 3).

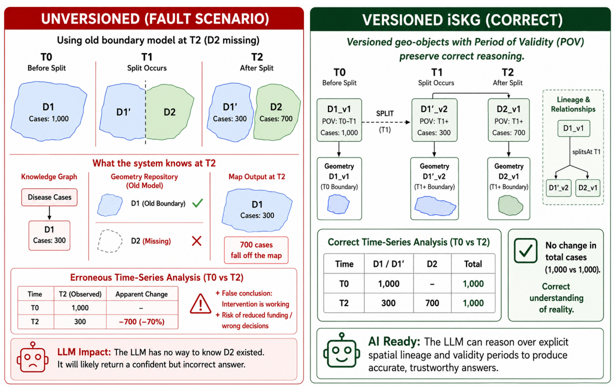

Figure 3. Temporal mismatch fault scenario

This is not merely a cartographic error. It is a semantic and temporal interoperability failure. The operational data and the boundary model are each valid within their own context, but they are not valid when combined without explicit temporal alignment and lineage. An LLM consuming the resulting data has no reliable way to infer that D2 exists, that D1 was split, or that the correct comparison requires aggregating D1′ and D2 at T2.

A versioned interoperable Spatial Knowledge Graph avoids this failure by preserving geo-object identity, lineage, geometry, and period of validity. The graph records that D1_v1 was valid from T0 to T1 and split at T1 into D1′_v2 and D2_v1, both valid from T1 onward. As a result, a Geo-GraphRAG workflow can retrieve the correct spatial lineage and determine that total cases remain 1,000 across the comparable geography. Rather than forcing the LLM to reconstruct this history implicitly, the iSKG makes the required spatial and temporal relationships explicit, queryable, and auditable.

3.4 Principles for Interoperable Spatial Knowledge Graphs

The preceding fault scenario illustrates why spatial knowledge graphs must be published with sufficient structure to support interpretation, validation, and reuse across organizational boundaries. Prior work identified the core principles required for a Spatial Knowledge Mesh, including decentralized ownership, publication of spatial knowledge graphs as reusable assets, dependency management, versioning, temporal validity, and interoperability through shared geographic references (McEachen and Lewis, 2023).

This paper builds on those principles by defining a metamodel for iSKGs. In this context, an iSKG must support stable geo-object identity; explicit semantic typing of geo-objects and relationships; shared geographic abstractions; temporal validity for geometries, attributes, and relationships; graph and ontology versioning; provenance and source authority; dependency management; change propagation; publication and discovery; and machine-actionable access for downstream systems and LLM-based workflows.

A key extension emphasized in this paper is the need for Integrity, Provenance, and Trust (IPT). For Geo-GraphRAG applications, it is not enough for an LLM to retrieve spatial relationships from a graph. The system must also be able to determine what the relationship means, who asserted it, when it was valid, which ontology version defined it, what source data supports it, and how it depends on other published knowledge. The metamodel presented in the next section operationalizes these requirements by treating geo-ontologies, geo-objects, relationships, provenance, temporal validity, and dependencies as first-order components of an interoperable spatial knowledge infrastructure.

4. Metamodel for Interoperable Spatial Knowledge Graphs

A central objective of the proposed metamodel is to provide organizations with a standardized way to publish iSKGs as governed knowledge assets. For an iSKG to participate in a Spatial Knowledge Mesh, it must expose enough semantic, temporal, and provenance structure for other systems to discover it, interpret it, validate it, and integrate it without relying on bespoke point-to-point agreements. The metamodel therefore defines not only how geo-objects and relationships are represented, but also how graphs, geo-ontologies, assertions, temporal validity, provenance, and dependencies are managed as first-order components of the knowledge infrastructure.

This structure is necessary because iSKGs are rarely isolated. A graph published by one organization may depend on a graph published by another, which may itself depend on a third. For example, Graph C may depend on Graph B, which depends on Graph A. If Graph A changes, the metamodel must provide enough information for that change to be detected, evaluated, and propagated transitively to dependent graphs. This requirement applies not only to instance graphs, but also to geo-ontologies. One organization may define a geo-ontology that extends or references another organization's geo-ontology, and may define relationship classes that connect geo-object classes governed by different authorities.

For example, a Ministry of Health may define a relationship class representing health facility catchment areas that connects health facilities governed by the ministry to villages maintained by another government authority. If the authoritative village graph changes because villages are split, merged, renamed, or reclassified, those changes should propagate to the health facility catchment iSKG so that downstream systems can identify which catchment relationships may need review. The purpose of the metamodel is to make these dependencies explicit, machine-actionable, and auditable, enabling independently governed spatial knowledge graphs to function as an interoperable mesh rather than as disconnected local graph databases.

4.1 Graphs as First-Order Knowledge Assets

The metamodel treats graphs, schema elements, concepts, geo-objects, and assertions as governed knowledge assets that can be identified, classified, versioned, attributed, and managed through a common structure. This reflects the principle that interoperability requires not only exchanging graph data, but also publishing the semantic structures that define how graph data should be interpreted. Ontologies, geo-ontologies, concept sets, and spatial knowledge graphs therefore evolve through time and must be represented as managed objects within the Spatial Knowledge Mesh.

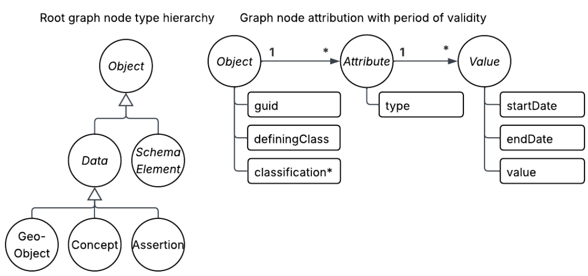

Figure 4. Geo-GraphRAG enables LLM to focus on reasoning

Figure 4 illustrates the root abstraction used by the metamodel. Circles represent graph node types, while rectangles represent attributes. The inheritance symbol is adapted from UML to indicate specialization among graph node types. At the root is the Object node type, from which both Data and Schema Element nodes inherit. Data nodes represent instance-level knowledge, including Geo-Objects, Concepts, and Assertions. Schema Element nodes represent definitional structures such as class definitions, relationship definitions, graph definitions, ontology definitions, and related schema components.

Following the CGR concept (Ebener, 2022), a geo-object represents a single identifiable geographic feature within the graph representation layer. Rather than duplicating the same real-world feature across multiple domain graphs, the geo-object provides a shared identity anchor to which names, classifications, geometries, attributes, and relationships can be attached over time. Assertions provide a mechanism for attaching domain-specific knowledge to shared geo-objects without treating every domain value as an intrinsic property of the feature. For example, a village may have different population values for disease intervention planning, taxation, census reporting, or disaster

https://doi.org/10.5194/isprs-archives-L-4-W1-2026-151-2026 | © Author(s) 2026. CC BY 4.0 License.

response. Modeling such values as assertions allows multiple domain-specific facts to reference the same geo-object while preserving source authority, provenance, and temporal validity.

All node objects share a common attribution pattern. Each Object has a globally unique identifier (guid), a reference to the class definition that defines it (definingClass), and an optional classification referencing a concept in an ontology or concept set. Each Object may have one or more Attribute nodes, and each Attribute may have zero or more Value nodes with a start date, end date, and literal value. This provides a uniform mechanism for representing attribute histories across geo-objects, concepts, assertions, and schema elements.

By applying the same object and attribution model to both data and schema elements, the metamodel allows the structures that define an iSKG to be governed in the same way as the instance data they describe. A geo-ontology can therefore be published, versioned, attributed, and referenced as a graph object, enabling downstream systems to determine what an object represents, which schema defined it, when its values were valid, and how it should be interpreted within a Spatial Knowledge Mesh.

4.2 Schema Elements for Interoperable Spatial Knowledge Graphs

The metamodel (Figure 5) distinguishes between instance data and the schema elements that define how that data may be structured, interpreted, and validated. Schema Element is the root abstraction for these definitional components. Whereas Data nodes (Figure 4) represent geo-objects, concepts, assertions, and relationships that exist within an instance graph, Schema Elements define the classes, graph types, attributes, and semantic structures used to govern those instances. This enables the schema itself to be represented, published, versioned, and referenced as graph-managed knowledge.

Object Class is the root abstraction for defining classes of graph objects. A class definition acts as a container for attribution: attributes are defined on classes and inherited by instances of those classes (Figure 6). Geo-Object Class specializes Object Class for geographic features. It is analogous to a feature class in GIS, but functions within the graph model as both a typing mechanism and an attribution container for geo-objects. A Geo-Object Class may define attributes such as name, identifier, geometry type, classification, or other properties required for a particular category of geographic feature.

Assertion Class is used to define domain-specific assertions that reference geo-objects. This allows organizations to attach analytical or domain-specific facts to shared geographic features without redefining the feature itself. For example, a health program, census authority, or emergency management agency may each assert different values or relationships involving the same village, depending on their mandate and data source. A Concept Class defines the attribution structure for concepts contained within a Concept Set, including enumerations, taxonomies, and ontologies.

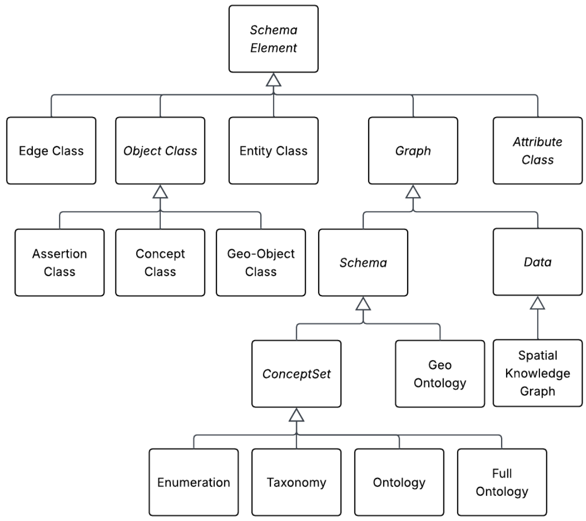

Figure 5. Metamodel Schema Elements

Edge Class defines the allowable relationship types between objects. These edges may connect geo-objects to other geo-objects, assertions to geo-objects, concepts to concepts, or schema elements to other schema elements (Figure 6). Edge Classes make relationship semantics explicit, allowing systems to distinguish between relationships such as containment, adjacency, dependency, service coverage, exposure, lineage, or ontology extension. This is essential for interoperability because an edge is only machine-actionable if its meaning is defined and discoverable.

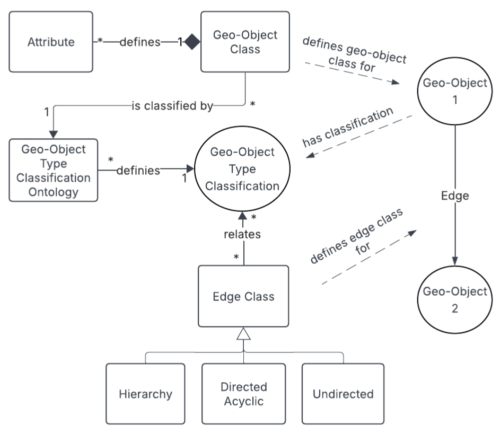

Figure 6. Attribute definition and edges between classes

A distinction is made between a Geo-Object Class and a Geo-Object Type. A Geo-Object Class is a structural schema abstraction that defines attributes, geometry expectations, identity rules, and relationship constraints for a category of geo-objects. A Geo-Object Type is a semantic classification that defines what kind of real-world feature the geo-object represents within a domain ontology. For example, Waterway may be defined as a Geo-Object Class with required attributes such as name, geometry, and flowDirection. Individual waterway geo-objects may then be classified using Geo-Object Types such as River, Canal, Stream, or Drainage Channel. In this example, the class defines the graph schema, while the type defines the domain meaning. This separation allows structural interoperability to be maintained while domain-specific classifications are governed through extensible ontologies published by different authorities.

Graph is the root abstraction for defining different types of graph assets (Figure 7). Schema graphs define semantic structures, while data graphs contain instance-level knowledge. Concept Set is a specialization of Schema used to define controlled semantic structures, ranging from simple enumerations to taxonomies and richer ontologies. Each Concept Set is associated with a Concept Class that defines the attribution of the concept nodes contained within it. Geo-Ontology is another specialization of Schema and defines a collection of Geo-Object Classes and Edge Classes used to construct interoperable spatial knowledge graphs. A Spatial Knowledge Graph is a specialization of Data and contains the geo-objects, assertions, relationships, and values that instantiate those schema definitions.

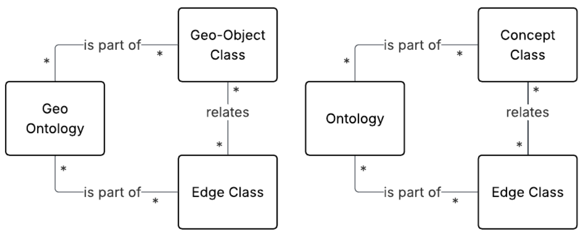

Figure 7. Ontology definitions

Entity Class provides a bridge between graph-based knowledge and relational representations. It defines tabular or flattened views derived from graphs for a specified period of time, allowing graph data to be materialized into relational structures when needed for analysis, exchange, reporting, or integration with legacy systems. This is important because many existing geospatial and administrative workflows continue to rely on tabular data products, even when the authoritative knowledge is maintained as a graph.

Together, these schema elements allow an iSKG to publish not only data, but also the machine-actionable definitions required to interpret that data. A consuming system can determine which geo-ontology defines a graph, which Geo-Object Classes define its nodes, which Edge Classes define its relationships, which Concept Sets provide controlled classifications, and which Attribute Classes define permitted values. This provides the foundation for interoperability, validation, provenance, temporal management, and change propagation across a Spatial Knowledge Mesh.

4.3 Temporal Validity and Versioning

Prior work introduced the Graph Repository as an abstraction for managing change over time in spatial knowledge graphs (McEachen and Lewis, 2023). A key distinction is that the repository manages periods of validity, not merely the time at which a record or edge was created, updated, or published. A period of validity describes the interval during which a graph object, relationship, attribute value, geometry, or schema element is understood to be valid in the real world.

This distinction allows the repository to preserve both the current understanding of the past and earlier understandings that existed at previous points in time. For example, a spatial knowledge graph describing health facilities, operational hierarchies, and administrative catchment areas may represent conditions in 2020. The original 2020 publication may later be found to contain errors and corrected in a subsequent version. Both versions are meaningful: the corrected version represents the current understanding of 2020, while the original version may be needed to audit operational decisions or reproduce analyses that were performed using the earlier data. The Graph Repository therefore supports versioned publications for the same period of validity, such as a 2020 v1 graph representing the original understanding and a 2020 v2 graph representing the corrected understanding.

The same principle applies to geo-ontologies. Spatial knowledge graphs change over time, but the schemas that define them also evolve. Earlier implementations treated geo-ontologies and SKGs as separate abstractions, limiting the ability to govern ontology change using the same temporal mechanisms applied to instance graphs. In the proposed metamodel, geo-ontologies and SKGs share the Graph abstraction, allowing both to be published, versioned, and managed by period of validity.

Figure 8 illustrates this pattern. In the graph repository, schema elements may have different periods of validity. For example, an administrative hierarchy may initially define a Province–District–Village structure for T0–T1, and later introduce a Ward level between District and Village for T1–T2. The repository can publish the geo-ontology that applies to each interval, allowing spatial knowledge graphs to be validated against the ontology version that was valid for the corresponding period. This ensures that instance graphs are interpreted relative to the correct schema, preserving temporal consistency, auditability, and trust across the Spatial Knowledge Mesh.

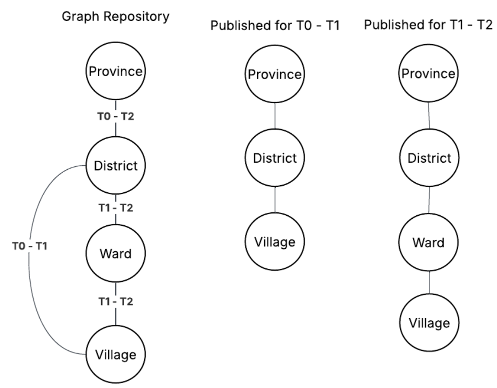

Figure 8. Geo-ontology periods of validity

4.4 Event-Based Graph Merging

The Graph Repository can be understood as maintaining the current "HEAD" state of the managed graph, analogous to a source-code repository. Although the repository stores objects, relationships, attributes, and schema elements across multiple periods of validity, it represents the current version of those historical periods. For production Geo-GraphRAG applications, however, querying the full temporal structure of the repository directly can be inefficient. A common pattern is therefore to materialize an iSKG for a specified period of validity into a labeled property graph or RDF triple store optimized for graph-query workflows.

A challenge arises when the source repository changes after a graph has been published. Additive changes can be appended, but updates to attribute values, geometries, or relationships may require existing graph elements to be replaced. Rebuilding an entire triple store or property graph after each change is inefficient, while incremental updates require knowing which objects, edges, and values are affected.

To address this, the metamodel supports event-based graph merging. Repository events record changes to objects, assertions, attributes, and edges within the Graph Repository. Published events are curated events derived from repository events for a specified publication, period of validity, and set of included object and edge types. Multiple repository events affecting the same object can be collapsed into a single published event representing the final state of that object at publication time. This allows downstream systems to consume official published changes rather than the full internal working history of the repository (Figure 9).

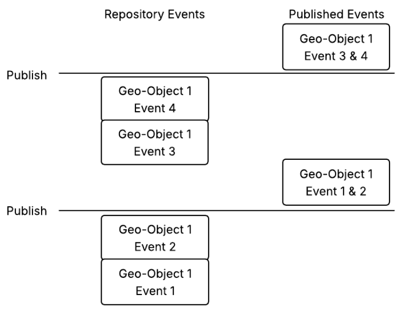

Figure 9. Collapsing change events for published graphs

Published events are grouped into commits. A commit identifies the publication version, period of validity, included object and edge types, metadata snapshot, and the repository events used to generate the publication. Downstream repositories can pull commits from remote sources, replay the published events as local repository events, and update local projections such as RDF triple stores or labeled property graphs. This allows materialized Geo-GraphRAG graphs to remain synchronized without requiring complete rebuilds.

4.5 Integrity, Provenance, and Trust

A key advantage of using iSKGs for Geo-GraphRAG is that model outputs can be made traceable to explicit graph evidence. When an LLM-based agent generates a graph database query, the query itself can be inspected, the returned graph objects and relationships can be examined, and the provenance of those objects can be followed. This provides a stronger basis for auditability than relying on unstructured context or unsupported model inference.

The objective of the iSKG provenance model is to preserve two related but distinct forms of provenance: the source graph and the data source. The source graph is the published iSKG from which a Geo-Object, Assertion, or relationship was retrieved. The data source identifies where the underlying values originated. This distinction is necessary because a graph publisher may transform, integrate, or republish data originally produced by another authority. For example, Organization A may publish an iSKG derived from spatial data tables produced by Organization B. A downstream Geo-GraphRAG workflow should be able to determine both that the retrieved object came from Organization A's published graph and that its underlying values originated from Organization B's source dataset.

Figure 10 illustrates this provenance structure. A Geo-Object is associated with a Data Source, which in turn is linked to a Source Authority representing the originating organization.

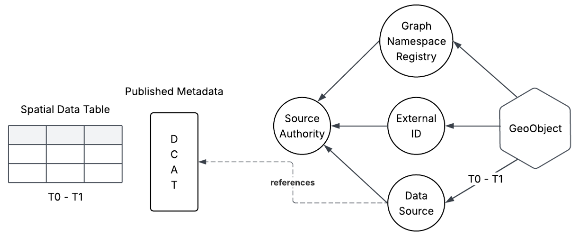

Figure 10. Graph and data provenance

The Geo-Object may also carry an External ID assigned by that source authority, while the Graph Namespace Registry supports the management of identifiers across published graphs. The Data Source may reference published metadata, such as DCAT, and may include information about the governance level of the source and the metadata profile used to describe it.

Governance Level	Description
Authoritative	Official source of record.
Official	Organizational publication.
Community Curated	Community-maintained dataset.
Research	Research-produced dataset.
Derived	Generated from source data.
Experimental	Prototype or pilot dataset.
Ad Hoc	Informally assembled dataset.

Table 1. Data source governance level

Tables 1 and 2 summarize the current governance level and metadata profile options used in the metamodel. These controlled values are intended to support consistent provenance interpretation across iSKGs and remain subject to refinement based on implementation experience, customer feedback, and ongoing work within the Open Geospatial Consortium.

Profile	Meaning
DCAT	General catalog metadata
GeoDCAT	Geospatial catalog metadata
ISO19115	Traditional geospatial metadata
STAC	EO/raster catalog metadata
SensorML	Sensor metadata
FHIR	Healthcare resource metadata
DataCite	Research metadata
Custom	Organization-specific structured metadata
AdHoc	Manual metadata without formal schema
None	No metadata available

Table 2. Data source metadata profile

To support federated identity management, each geo-object is assigned a globally unique identifier composed of a namespace and a locally unique identifier. The namespace identifies the authority or GeoPrism instance responsible for issuing and resolving the identifier, while the local identifier distinguishes the geo-object within that namespace. Because the architecture is federated, namespaces must be governed in a way that prevents collisions without requiring a single centralized registry. This can be achieved through globally delegated namespace mechanisms, such as DNS-based identifiers, persistent URI patterns, or formally registered namespace records. External identifiers from source organizations are maintained separately from Geo-Object GUIDs, allowing users to reference familiar local identifiers while preserving a stable graph identity across merged iSKGs.

https://doi.org/10.5194/isprs-archives-L-4-W1-2026-151-2026 | © Author(s) 2026. CC BY 4.0 License.

5. Implementing a Geo-GraphRAG

The Open Geospatial Consortium AI-DGGS Pilot and related work with the U.S. Army Corps of Engineers demonstrated that spatial knowledge graphs can provide effective grounding infrastructure for Geo-GraphRAG workflows (Open Geospatial Consortium, 2025). In these implementations, geoinformatics data were represented as graph relationships that an LLM-based agent could query directly, rather than requiring the model to infer spatial context from unstructured text or disconnected map layers.

A typical Geo-GraphRAG workflow involves three steps. First, relevant geoinformatics data are transformed into a spatial knowledge graph, such as RDF triples or a labelled property graph. Second, the agent is configured with the geo-ontology that defines the available geo-object classes, relationship classes, and attributes. This provides the model with the semantic context needed to understand what can be queried. Third, the model translates natural-language questions into graph queries, such as GeoSPARQL or Cypher, retrieves explicit graph evidence, and uses that evidence to generate an answer.

Current implementations have demonstrated graph-query generation using commercial LLMs for both GeoSPARQL and Cypher. However, open models tested to date have not consistently generated graph queries of the required complexity. This limitation is expected to diminish as open models improve, but it reinforces the value of reducing the reasoning burden placed on the model through explicit geo-ontologies and well-structured graph access patterns.

Open-source reference implementations are available to demonstrate how Geo-GraphRAG can be incorporated into user-facing applications. The open-source Kaleidoscope application provides a chatbot-oriented interface for interacting with geospatial knowledge infrastructure and auditing model responses through maps and graph exploration. A related DGGS-focused prototype demonstrates how Geo-GraphRAG workflows can interact with DGGS-based services[1].

6. Discussion and Future Work

Although the original metamodel described in McEachen and Lewis (2023) did not support a federated architecture, it proved useful for constructing spatial knowledge graphs that support Geo-GraphRAGs, as demonstrated through projects with the U.S. Army Corps of Engineers and collaborative research within the OGC (Open Geospatial Consortium, 2024; Open Geospatial Consortium, 2025). This paper extends that work by introducing a metamodel designed to support federated implementation of interoperable Spatial Knowledge Graphs (iSKGs) as the foundation for a Spatial Knowledge Mesh.

Several aspects of the design remain under development. One area of interest is alignment with OGC Building Blocks, which provide modular, reusable specifications for promoting interoperability across geospatial systems. The objective is to express geo-ontologies in terms of OGC Building Blocks to facilitate integration with external tools and workflows while supporting change propagation and graph interoperability.

Another important area concerns versioning. Geo-ontologies evolve through changes in semantics and can therefore employ conventional semantic versioning. Geo-Objects, however, represent real-world entities whose attributes, geometries, and relationships change through time. As a result, Geo-Object versions must be associated with representative dates rather than publication order, allowing historical states to be added retrospectively while preserving temporal consistency. Following review by members of the OGC community, the proposed metamodel will be implemented in GeoPrism[2]. Success will ultimately be measured by how easily organizations can independently create, publish, and maintain interoperable Spatial Knowledge Graphs. The long-term vision is a Spatial Knowledge Mesh in which authoritative graph content can be published, discovered, merged, and maintained across organizational boundaries while preserving integrity, provenance, trust, and temporal consistency.

References

Ebener, S., 2022. HIS geo-enabling: Guidance on the establishment of a common geo-registry for the simultaneous hosting, maintenance, update and sharing of lists as well as associated hierarchies and spatial data. Version 2.0. Health GeoLab Collaborative. healthgeolab.net/DOCUMENTS/Guidance_Common_Geo-registry_Ve2.pdf, last access: 14 June 2026.

Edge, D., Trinh, H., Cheng, N., Bradley, J., Chao, A., Mody, A., Truitt, S., Metropolitansky, D., Ness, R. O., Larson, J., 2024. From Local to Global: A GraphRAG Approach to Query-Focused Summarization. arXiv preprint arXiv:2404.16130. doi.org/10.48550/arXiv.2404.16130.

McEachen, N., Lewis, J., 2023. Enabling Knowledge Sharing by Managing Dependencies and Interoperability Between Interlinked Spatial Knowledge Graphs. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLVIII-4/W7-2023, 117–124. doi.org/10.5194/isprs-archives-XLVIII-4-W7-2023-117-2023.

Open Geospatial Consortium, 2024. OGC Climate and Disaster Resilience Pilot 2024.2 Engineering Report. OGC Public Engineering Report 25-010. doi.org/10.17605/OSF.IO/CE34A.

Open Geospatial Consortium, 2025. AI-DGGS Pilot for Disaster Management: Final Engineering Report. ai-dggs-c43a9e.pages.ogc.org/documents/D001/document.html#annex-terraframe

Pan, S., Luo, L., Wang, Y., Chen, C., Wang, J., Wu, X., 2024. Unifying Large Language Models and Knowledge Graphs: A Roadmap. IEEE Transactions on Knowledge and Data Engineering. doi.org/10.1109/TKDE.2024.3352100.

Sikka, V., Sikka, V., 2025. Hallucination Stations: On Some Basic Limitations of Transformer-Based Language Models. arXiv preprint arXiv:2507.07505. doi.org/10.48550/arXiv.2507.07505.

Song, P., Han, P., Goodman, N., 2026. Large Language Model Reasoning Failures. Transactions on Machine Learning Research. arXiv:2602.06176. doi.org/10.48550/arXiv.2602.06176.

[1] github.com/terraframe/kaleidoscope, github.com/terraframe/kaleidoscope-dggs

[2] github.com/terraframe/geoprism