

## Article

# Smart City Ontology Framework for Urban Data Integration and Application

Xiaolong He <sup>1</sup>, Xi Kuai <sup>1,2,\*</sup> , Xinyue Li <sup>1</sup>, Zihao Qiu <sup>1</sup>, Biao He <sup>1,2</sup> and Renzhong Guo <sup>1,2</sup>

<sup>1</sup> Research Institute for Smart Cities, School of Architecture and Urban Planning, Shenzhen University, Shenzhen 518060, China; 2210324001@email.szu.edu.cn (X.H.)

<sup>2</sup> MNR Key Laboratory for Geo-Environmental Monitoring of Great Bay Area, Shenzhen University, Shenzhen 518060, China

\* Correspondence: kuaixi@szu.edu.cn

## Highlights

### What are the main findings?

- A hierarchical ontology (SMOF) with universal and extended properties and a concise relation scheme that draws on authoritative standards/ontologies (e.g., IFC, CityGML, SSN/SOSA) to support city-wide, cross-domain data integration.
- Combined quantitative analyses, LLM as judge assessment, expert evaluation, and two empirical scenarios confirm SMOF's structural soundness, conceptual richness, and capacity to integrate heterogeneous data for querying and reasoning.

### What is the implication of the main finding?

- By harmonizing heterogeneous data and semantics, SMOF enables coordinated urban services ranging from emergency management to transportation and infrastructure.
- Its scalability and reusability provide a foundation for extending ontology-driven approaches to broader domains of smart city governance and decision-making.

## Abstract

Rapid urbanization and the proliferation of heterogeneous urban data have intensified the challenges of semantic interoperability and integrated urban governance. To address this, we propose the Smart City Ontology Framework (SMOF), a standards-driven ontology that unifies Building Information Modeling (BIM), Geographic Information Systems (GIS), Internet of Things (IoT), and relational data. SMOF organizes five core modules and eleven major entity categories, with universal and extensible attributes and relations to support cross-domain data integration. SMOF was developed through competency questions, authoritative knowledge sources, and explicit design principles, ensuring methodological rigor and alignment with real governance needs. Its evaluation combined three complementary approaches against baseline models: quantitative metrics demonstrated higher attribute richness and balanced hierarchy; LLM as judge assessments confirmed conceptual completeness, consistency, and scalability; and expert scoring highlighted superior scenario fitness and clarity. Together, these results indicate that SMOF achieves both structural soundness and practical adaptability. Beyond structural evaluation, SMOF was validated in two representative urban service scenarios, demonstrating its capacity to integrate heterogeneous data, support graph-based querying and enable ontology-driven reasoning. In sum, SMOF offers a robust and scalable solution for semantic data integration, advancing smart city governance and decision-making efficiency.



Academic Editors: Guillermo del Campo Jiménez, Asun Santamaría and Pierluigi Siano

Received: 13 August 2025

Revised: 16 September 2025

Accepted: 30 September 2025

Published: 3 October 2025

**Citation:** He, X.; Kuai, X.; Li, X.; Qiu, Z.; He, B.; Guo, R. Smart City Ontology Framework for Urban Data Integration and Application. *Smart Cities* **2025**, *8*, 165. <https://doi.org/10.3390/smartcities8050165>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** ontology framework; smart city; multi-source data integration; knowledge graph

## 1. Introduction

As reported by the United Nations Population Fund (UNFPA) [1], urban residents already account for 56% of the global population, around 4.4 billion people, and this share is projected to approach 70% by 2050. Rapid urbanization has therefore magnified the complexity of city governance. At the same time, advances in information and communication technologies (ICT) have enabled every urban department or agency to build dedicated datasets for its own objectives. The architecture, engineering, and construction industry adopts BIM for project life-cycle data; the geomatics community collects, analyzes, and visualizes spatial data through GIS; and relational and time-series databases record transactions and sensor readings across municipal departments. The resulting surge in data volume and variety has catalyzed the emergence of the smart city paradigm, which seeks to strategically harness ICT and diverse urban datasets to address governance complexities, enhance sustainability, and improve citizens' quality of life [2–4].

Within this smart city paradigm, the combined use of BIM/GIS/IoT data flows and other representations of the physical world equips stakeholders with insights from macro-scale urban planning to micro-scale sensor behavior. Together, these data enable a comprehensive understanding of city operations [5]. Multi-source heterogeneous urban data have thus become indispensable for public service management, resource optimization, and resilience enhancement. Yet, the multidisciplinary nature of urban science poses formidable challenges [6]. First, the diversity of data representation and granularity complicates integration. Different standards adopt distinct modeling schemes and levels of detail. For instance, CityGML encodes city-scale geometries in XML, whereas the IFC standard captures fine-grained building components and topology in STEP format. As a result, mismatches in scale and structure hinder seamless data fusion. Second, semantic ambiguity across institutions hampers the effective use of data. Different agencies often generate overlapping datasets that describe the same urban phenomenon but adopt inconsistent labeling and classification [2,3]. For example, one dataset may record “urban greening,” while another refers to “urban parks.” This inconsistency leads to semantic fragmentation and significant challenges for interoperability. Third, many application scenarios demand simultaneous integration of heterogeneous data types. Complex urban services often require multiple datasets to be combined in real time. For instance, a fire emergency response typically calls for the seamless integration of BIM/GIS/IoT and relational data. The absence of effective integration mechanisms in such cases seriously hampers the timely application of heterogeneous urban data [7,8].

These challenges, including diverse data representations and granularities, semantic ambiguities, and the need for real-time integration, underscore the urgent need for robust mechanisms for semantic interoperability. To address these integration hurdles, ontologies and knowledge graphs are considered effective means [9]. An ontology provides a formalized specification of domain concepts, attributes, relationships, and constraints, offering a unified language for applications hindered by structural and semantic disparities. A knowledge graph instantiates this ontology by mapping and extracting core semantics from heterogeneous sources into a machine-readable network, thereby enabling integrated querying and reasoning across diverse datasets [10]. In this context, a smart city ontology can be defined as a formal, structured representation of concepts, attributes, and their interrelationships within the massive knowledge sources of a city. It is designed to enable semantic interoperability between different urban data sources for specific purposes [3].

Previous research has developed ontologies either for particular business processes—such as flooding or fire safety—or for single data standards such as IFC or CityGML. However, these efforts usually focus solely on their specific mandate and do not treat city-wide entities as reusable management units, nor do they impose a layered structure on those entities. More recent attempts have led to representative frameworks such as UrbanKG (UKG), KM4City, and CIMO. While these models demonstrate the potential of ontology-driven integration—covering multimodal entities, large-scale municipal datasets, or BIM/GIS/IoT extensions—they remain limited in scope. Specifically, they are often designed for particular application domains, lack comprehensive coverage of heterogeneous city assets, and provide only partial support for extensibility across layers of urban knowledge. As a result, existing urban ontologies, and the knowledge graphs built on top of them, frequently lack reusability and extensibility, and few attempt to break data silos and model urban assets hierarchically from a holistic perspective.

Building on these gaps, our research is guided by the following questions:

RQ1: How can multi-source heterogeneous urban data be semantically integrated to address issues of semantic fragmentation and representational differences?

RQ2: How can an ontology framework be constructed to encompass the vast range of urban management objects and associated knowledge?

RQ3: How can the proposed ontology framework be applied and validated in real urban scenarios?

To answer these questions, we propose the SMOF, designed to support integration and application of heterogeneous urban data from a city-wide management viewpoint [11]. The main contributions are threefold: (1) We formulate competency questions, clarify foundational design principles, and collect authoritative City Information Modeling (CIM) standards as knowledge sources. (2) We construct the SMOF's hierarchical entity taxonomy, universal attributes, and relationship patterns, and detail its mappings to existing ontologies. (3) We answer the competency questions using SPARQL-OWL (SPARQL for OWL, where SPARQL denotes the Simple Protocol and RDF Query Language, and OWL refers to the Web Ontology Language) queries. We then evaluate SMOF through quantitative metrics, expert scoring, and an LLM as judge protocol, and demonstrate its practical utility in two scenarios.

The remainder of the paper is organized as follows. Section 2 reviews related work on multi-source urban data integration. Section 3 presents the workflow and design of the SMOF in detail. Section 4 reports multi-dimensional evaluation results, including quantitative metrics, LLM as judge analysis, expert-based assessment, and empirical validation in two scenarios: fire emergency and traffic congestion scenarios. The paper's findings are detailed in Section 5, with concluding remarks in Section 6.

## 2. Related Work

### 2.1. Integration and Application of Multi-Source Heterogeneous Urban Data

Integrating and exploiting heterogeneous urban data from multiple sources has become a focal research topic under the data-driven smart city paradigm. Scholars generally follow two technical routes—deep learning pipelines and knowledge graph pipelines—to derive actionable insights from voluminous, structurally diverse datasets.

(1) Deep-learning-centered studies. Recent studies use deep neural networks to extract latent representations from multimodal urban data and support a wide range of analytics. UrbanVLP [12] and RSGPT [13] exemplify this trend by coupling visual streams with text to pre-train or align models for downstream scene understanding. Adegun et al. (2024) combine YOLOv8 with an ontology guided pipeline to detect buildings, roads, and coastlines, and then populate a structured knowledge representation that can be

queried via SPARQL for semantic post-analysis [14]. Zou et al. and Alahi et al. review cross-domain data fusion techniques for urban analytics and categorize deep learning methods into feature-, alignment-, contrast-, and generation-based fusion; they further map urban application scenarios to planning, economy, society, energy, transport, public safety, and environment [2,15]. In recent studies, Choi et al. [16] propose an AI multi-agent data integration framework for urban building energy modeling based on LLMs, enabling the automated integration and analysis of GIS and relational data related to urban buildings. However, research on LLM based data integration remains at an early stage and still requires further expansion of multi-source data capabilities and the development of ontologies to enhance adaptability.

Despite their power, deep models suffer from well-known drawbacks. Their decision logic is often opaque, which hampers adoption in urban service contexts where explainability is mandatory [17]. Moreover, training effective models typically demands large labeled corpora, yet annotating city-scale data is costly and error prone, injecting noise that degrades generalization.

(2) Knowledge-graph-centered studies. Knowledge graphs are widely applied in smart city research, yet their semantic validity ultimately depends on well-defined ontologies that provide the conceptual vocabulary and design rules. Guided by ontologies, knowledge graphs represent entities and relations explicitly, ensuring unified semantics for targeted applications. Representative efforts span infrastructure asset-management workflows [18], multi-domain information integration for flood-risk assessment [19], and POI graphs that fuse spatial/semantic links to support personalized routing [20]. Panagiotopoulou et al. [3] address the problem of fragmentation and insufficient integration of multi-dimensional smart city performance indicators (sustainability, resilience, inclusiveness) by embedding these indicators into an ontology, thereby proposing an ontology oriented framework for the integration of performance assessment indicators. Liu et al. survey smart city data integration solutions and conclude that ontologies and knowledge graphs offer superior interpretability and less reliance on labeled data than deep-learning-only pipelines—advantages that are crucial for disaster management, environmental monitoring, and other semantics intensive scenarios [21].

The value of a knowledge graph fundamentally depends on its underlying ontology; therefore, numerous researchers have sought to construct ontology models adaptable to multi-source data or complex scenarios by leveraging reference data standards, competency questions, and fundamental design principles. Huang et al. [22] integrate three Open Geospatial Consortium standards—CityGML, IndoorGML, and the SensorThings API—into a single knowledge graph that powers smart home automation, e-health, and fire evacuation systems. Building on the CityRDF ontology, Ding [23] demonstrate semantic interoperability for CityGML and assembled a knowledge graph that supports representation and querying of diverse 3D geospatial datasets. Grisiute et al. [24] constrain the scope and objectives of the ontology by employing competency questions, thereby providing an ontology model capable of automatically generating three-dimensional master plans. Tok et al. [25] review key considerations regarding ontology design principles identified in comprehensive surveys and apply them to his own ontology. This effort further ensured the rigor and methodological soundness of his ontology construction process for smart city infrastructure.

In summary, knowledge-graph-based approaches can effectively integrate multi-source heterogeneous urban data, while the incorporation of competency questions, fundamental ontology design principles, and reference knowledge standards ensures that the underlying ontologies remain rigorous and consistent. In turn, these ontologies provide

the semantic backbone that allows knowledge graphs to function as robust, interoperable platforms for data storage, querying, and reasoning in urban services.

## 2.2. Ontology Construction for Multi-Source Urban Data

Urban ontologies, the semantic foundation for knowledge-graph construction, aim to unify fragmented urban concepts, attributes, and relations through formal modeling. Early studies highlighted the integrative potential of knowledge-graph technologies to organize diverse urban business processes. For example, Abid et al. [26] leverage DBpedia to link issue reports, sensor observations, and administrative elements, enabling semantic linking and interoperability of municipal services. Sujata et al. [27] propose the SMELTS framework spanning the social, management, economic, legal, technological, and sustainability dimensions; although not a strict ontology, it provides valuable guidance for the hierarchical construction of urban knowledge systems. The Smart City Services Ontology (SCSO) [28] adopts a service-oriented perspective to develop a layered ontology organized around the policy–service–user hierarchy, supporting applications that involve municipal management, citizen interactions, and IoT devices. Notably, SCSO does not explicitly model the full range of urban entities; instead, it employs a generic framework complemented by extendable domain ontologies, ensuring adaptability across diverse scenarios.

Despite the advances above, most urban ontologies remain focused on specific business objects, lack structured organization, and fail to represent latent objects of urban governance; moreover, business-driven attributes and relations are difficult to reuse across scenarios [29,30]. Recent studies address these gaps via hierarchical modeling with scenario-specific extensions. Pereira et al. [31] adopt a BIM-inspired modular design to integrate knowledge of the built environment, water bodies, infrastructure and services, and mobility, supporting complex decision-making in planning and construction. Xu et al.'s work on CIM semantic trees [32] shows that graded, hierarchical schemes improve the organization and analysis of urban entities. Although their study did not incorporate other spatiotemporal data sources such as GIS and IoT, it demonstrated the effectiveness of hierarchical structures in semantic analysis and knowledge organization. Drawing on the SCSO idea, CIMO separates BIM/GIS/IoT ontology construction into a generic core and an extension layer, building the foundational ontology at the generic level and refining it for concrete scenarios in the extension layer [33]. Taken together, although they do not yet adopt a city-wide perspective that integrates additional spatiotemporal sources, these studies advance a practical recipe for complex urban knowledge architectures and cross-domain extensibility: organize entity models in a modular, hierarchy-first manner, establish a generic ontology, then extend it for specific scenarios. This approach connects disparate domains while maximizing extensibility and reducing complexity.

Based on the research questions and the current state of related work, this study proposes the following hypotheses:

**Hypothesis 1.** *By employing ontology methodologies in conjunction with competency questions, ontology design principles, and domain standards, it is possible to effectively mitigate semantic fragmentation and representational inconsistencies, thereby enabling the unified semantic representation of multi-source heterogeneous urban data.*

**Hypothesis 2.** *Adopting a hierarchical entity modeling approach, while balancing generality and extensibility in the design of attributes and relations, can achieve systematic modeling of diverse urban objects and their associated knowledge.*



**Hypothesis 3.** *Through empirical applications in representative smart city scenarios, the proposed ontology framework can be validated in terms of effectiveness and applicability, thereby demonstrating its practical value in real urban environments.*

To anchor subsequent comparisons, and following the representative directions identified by Wang [34], we adopt three widely cited frameworks as baselines: UrbanKG [35], an aggregated city knowledge graph built on ontology (e.g., POIs, regions, users, brands, images) designed to bridge modality and structural gaps for graph-based analytics; KM4City [36], a municipal semantic backbone integrating administration, road networks, POIs, public transport, sensors, temporal aspects, and meta-information to support city services at scale; CIMO [33], a BIM/GIS/IoT ontology that, in line with the core-plus-extensions idea noted above, layers a generic core with scenario-level refinements to obtain richer built-environment semantics. These baselines span complementary emphases—data aggregation (UrbanKG), service-centric municipal integration (KM4City), and layered BIM/GIS/IoT semantics (CIMO). We evaluate SMOF against them using objective metrics and a mixed subjective protocol, with results in Sections 4.2 and 4.3.

### 3. Ontology Framework Design

#### 3.1. Overall Workflow

In this paper, the ontology framework denotes the complete package that integrates the ontology-engineering methodology and the schema specification (classes, attributes, and relations). It also includes the application workflow that operationalizes the schema through queries and reasoning. Entity modules are the top-level domain partitions in the schema; the entity hierarchy is the class taxonomy within each module; and the knowledge graph is the instance-level dataset instantiated from the ontology schema.

Developing a city-scale ontology is a demanding engineering task. We therefore adopt Methontology, the life cycle methodology devised by the AI Lab of the Technical University of Madrid, whose phases—specification, conceptualization, formalization, integration, implementation, and evaluation/maintenance—provide a proven scaffold for rigorous ontology work [37]. To meet the realities of heterogeneous, multi-source urban data, we adapt Methontology into a three-stage workflow (Figure 1): (i) Specification and knowledge acquisition—define competency questions (CQs), gather authoritative knowledge from standards and mature ontologies, and set foundational design principles, ensuring clarity, consistency, and scope compliance; (ii) Conceptualization and fusion—adopt a top-down strategy to build hierarchical core entity modules together with universal and extended attributes and relations, balancing generality and domain specificity; and (iii) Implementation and evaluation—instantiate the ontology in Protégé, verify CQs via SPARQL-OWL queries, and assess performance through expert review and scenario-based validation. This adaptation preserves the rigor of Methontology while directly addressing the integration and reasoning challenges posed by city-scale, cross-domain datasets.

#### 3.2. Specification and Knowledge Foundations

The scientific validity and operational value of an ontology depend not only on methodological rigor but also on explicit specification and systematic knowledge acquisition. In the workflow outlined in Section 3.1, this stage involves three essential tasks: (1) competency questions, which anchor the ontology to concrete requirement directions; (2) knowledge sources, which ensure alignment with authoritative standards and mature ontologies; and (3) modeling principles, which regulate granularity, naming, and classification. These tasks correspond to the specification phase of Methontology [37] and are widely acknowledged in ontology engineering practice [23–25]. Their purpose is to provide

methodological transparency and to ground the design process in verifiable requirements, credible references, and disciplined modeling standards.

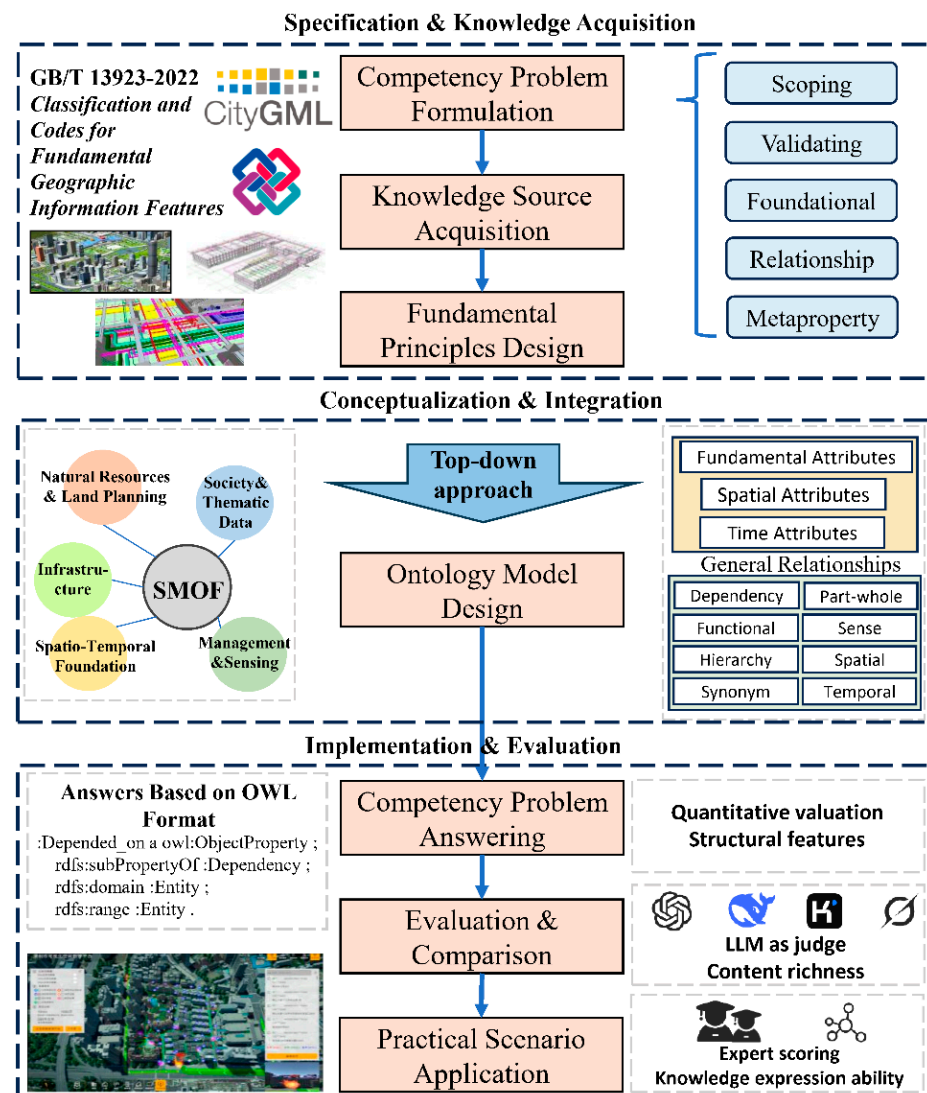


Figure 1. Overall workflow design.

### 3.2.1. Competency Questions

Competency questions specify the requirements that must be addressed during ontology development. They are typically defined by domain experts prior to ontology construction and, once the ontology is built, are answered by developers using ontology query languages (commonly SPARQL). This process ensures that the development remains aligned with the original objectives. Drawing on Keet's typology [38], the CQs were categorized into five groups: scope delimitation, verification, foundational alignment, relation-oriented, and meta-attributes. To adapt this framework to the context of this study, the CQs (as described in Table 1) were generated and refined by domain experts, ensuring that the general structure was preserved while being extended to meet the practical requirements of smart cities. For instance, within the verification CQs, we assessed SMOF's cross-domain mappings to widely used urban ontologies and required that it encode the spatial information needed to integrate heterogeneous multi-source data. In the relation-oriented CQs, CQ7 checks whether core relations (part-whole, dependency) are correctly defined, and CQ8 tests logical coherence via transitivity and reasoning. Contradictory answers would signal internal inconsistencies. These adaptations align

the CQs with multi-source integration needs, guiding development and later serving as validation instruments.

**Table 1.** Competency question categories and examples.

CQ Category	Purpose	Concrete Question
Scope-defining	Delimit the thematic range of the ontology	CQ1: Can SMOF represent hierarchical structures among entities, the intrinsic information of each entity, and inter-entity relations?
		CQ2: Which urban domains can SMOF cover?
Verification	Validate ontology content	CQ3: Can SMOF encode spatial information in terms of address, latitude–longitude, and topology?
		CQ4: Can SMOF map to the macro-classes defined in KM4City, UrbanKG, and related ontologies?
Foundational alignment	Align domain entities with foundational ontologies	CQ5: Can SMOF interoperate with ontologies that capture spatial and geometric information?
		CQ6: Can SMOF map to sensor-oriented ontologies?
Relation-oriented	Characterize key relational patterns	CQ7: Can SMOF express basic relations such as whole-part and dependency?
		CQ8: Can SMOF support semantic reasoning based on the defined relations?
Meta-attribute	Specify essential attributes	CQ9: Can SMOF represent common attributes such as name and state?
		CQ10: Can SMOF ensure global identity uniqueness for entities via meta-attributes?

### 3.2.2. Knowledge Sources

Knowledge of urban services comes from a variety of sources, such as standards/specifications, scholarly and professional publications, reports, and news media [39]. Among them, standards and specifications are particularly crucial as they provide systematic, standardized, and professional knowledge. Even though urban data standards typically do not specify ontologies for data objects, their conceptual representations of domain objects can underly the creation of urban entities. At the same time, there are numerous ontologies in the urban domain, and reusing existing authoritative ontologies in new ontologies is also an essential part [30].

Based on these considerations, we built the ontology on authoritative standards and ontologies, integrating nine national and international standards and five reference ontologies spanning urban governance, three-dimensional modeling, and component/asset management. A summary of representative standards and ontologies is provided in Table 2, while the complete list with detailed specifications and references is available on figshare. We prioritized widely recognized, practice-tested standards—CityGML, IFC, and SSN/SOSA—which offer comprehensive conceptual models and are broadly applied in smart-city projects, providing a robust semantic foundation. National specifications



(GB/T 40765-2021, GB/T 36625.5-2019) ensure alignment with China’s urban-governance requirements, yielding both international interoperability and local applicability. These standards and ontologies can be accessed through official platforms, such as the National Public Service Platform for Standards Information (<https://std.samr.gov.cn/gb/>), accessed on 1 October 2025, for Chinese specifications, and through corresponding websites and publications for international standards like CityGML and IFC.

**Table 2.** Summary of representative standards and ontologies incorporated into the SMOF.

Category	Representative Standards/Ontologies	Purpose/Role in Ontology Construction
Geospatial Standards	GB/T 40765-2021 [40]; GB/T 13923-2022 [41]; CityGML [42]; CJJ/T 197-2018 [43]	Define geographic ontology models, classification codes, and 3D city object schemas
Urban Management	GB/T 30428.2-2013 [44]; GB/T 28590-2012 [45]; GB/T 36625.5-2019 [46]	Provide taxonomies and coding for managed components, underground facilities, and municipal infrastructure data
Building Information	IFC 4x1 [47]; GB/T 51269-2017 [48]	Standardize BIM information structure and coding
Cross-domain Ontologies	KM4City [36]; UrbanKG [35]; SSN/SOSA [49]	Integrate multi-domain municipal, transport, and POI datasets
Temporal & Semantic Standards	TimeOWL [50]; GB/T 32853-2016 [51]	Support temporal reasoning and unified geospatial classification

### 3.2.3. Modeling Principles

Given the diversity of knowledge and uneven granularity across urban domains, explicit principles are still required to guide the design of the SMOF. Drawing on prior studies, we distilled the following modeling principles for the SMOF [25,52]:

(1) Modeling scope. The SMOF focuses on objectively existing urban entities that carry clear management imperatives, including natural features, manmade infrastructure, and socio-economic objects, all of which exhibit spatio-temporal attributes and rich semantics. The SMOF inherits existing classification systems to ensure compatibility. Pragmatic extensions are then added to meet data-exchange requirements in urban management.

(2) Modeling scheme. The ontology adopts a four-dimensional framework [52]—Entities (*E*), Properties (*P*), Relations (*R*), Instances (*I*)—formally,

$$SMOF = (E, P, R, I) \quad (1)$$

where entities are urban objects such as buildings, sensors, or natural resources; properties capture intrinsic and extrinsic features (such as location, ID and state); relations encode interactions like spatial, compositional and functional; instances are concrete manifestations generated via multi-source mapping.

Urban management objects are numerous, domain-specific, and strongly hierarchical. To address this, we employ a linear classification strategy [32,53,54]. Entities are partitioned by attributes and features into discrete modules and successive levels, ultimately yielding a

hierarchy-structured taxonomy. Each class at level  $k$  is a specialization of its parent at level  $k - 1$ , and classes within the same level are mutually exclusive. Formally,

$$E = \cup_{k=1}^n E^k, E^k = \cup_{i=1}^{m_k} E_i^k \quad (2)$$

where  $k = 1, 2, \dots, n$  denotes the hierarchy level,  $E^k$  the set of entities at level  $k$ ,  $m_k$  the number of classes in that level, and  $E_i^k$  the  $i$ -th class, such that

$$E_i^k \in E_{parent}^k, E_i^k \cap E_j^k = \emptyset \quad i \neq j \quad (3)$$

where  $E_{parent}^{k-1}$  represents the parent class set of level  $k$ .

### (3) Classification criteria.

(a) Entities, attributes, and relations must strictly follow the reference sources. Both intension and extension are defined unambiguously to avoid vagueness or polysemy, thereby ensuring accurate and consistent information exchange.

(b) A unified prefix–namespace scheme underpins all identifiers. Reserved extension slots allow scenario-driven enlargement and explicit mappings to external ontologies, fostering semantic interoperability and data linkage.

(c) When semantic conflicts arise between standards, the overriding objective is loss-free data exchange. Modeling choices therefore prioritize representations that preserve all critical elements during transfer and transformation.

(d) The ontology framework is designed for continuous iteration. Categories, attribute definitions, and relation sets will be refined as management requirements evolve, technologies advance, and empirical feedback is obtained from deployment.

(e) To secure a common semantic core while retaining domain specificity, attributes are partitioned into universal and extended sets. Relations are likewise divided into universal and extended groups, which strikes a balance between universality and situational fitness [28,33].

By anchoring the SMOF on these foundations, we provide a consistent yet extensible framework that can reconcile redundant standards, integrate core data models, and flexibly serve diverse smart city services.

### 3.3. Core Entity Module Design

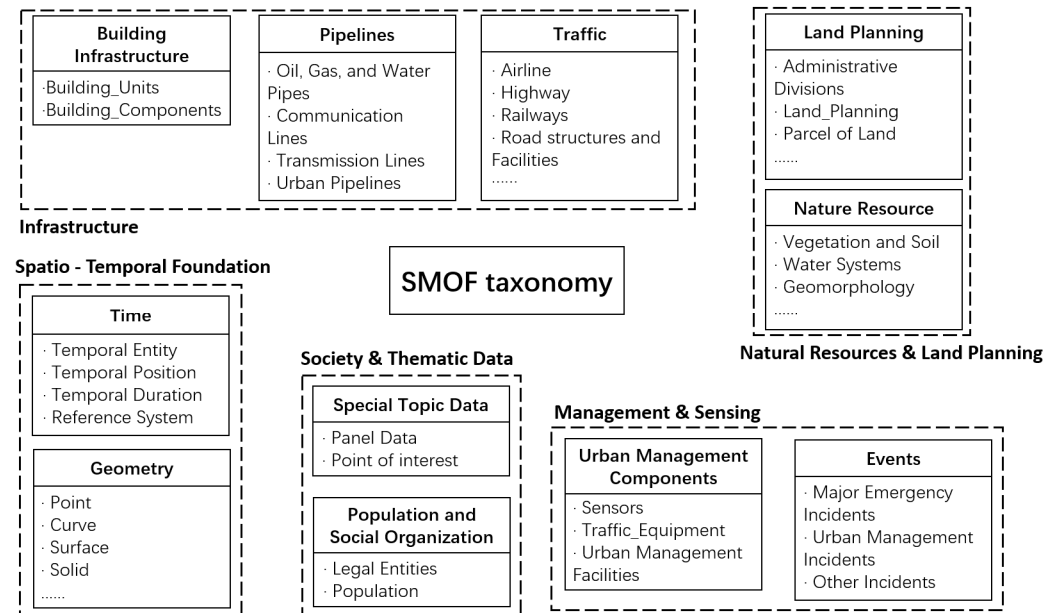
Following the principles and knowledge bases outlined in Section 3.2, we developed the SMOF by extracting classifications, attributes, and relations from each reference standard, semantically consolidating overlapping content, and disambiguating redundancies. The resulting entity hierarchy constitutes the backbone that links data integration, semantic definition, scenario extension, and decision support. Guided by the source standards, the SMOF partitions urban entities into five core modules, eleven primary classes, and forty-two intermediate classes. The modules—Land Planning, Natural Resources, Building Infrastructure, Pipelines, Traffic, Time, Geometry, Population and Social Organizations, Special Topic Data, Urban Management Facilities and Components, and Events—are organized into the conceptual hierarchy illustrated in Figure 2.

**Infrastructure module.** We treat infrastructure as the union of Buildings, Pipelines, and Transport, drawing chiefly on GB/T 28590-2012, the IFC 4×1 ontology, CJJ/T 197-2018, and GB/T 32853-2016 for tiered classification.

**Natural resources and planning entities.** The primary classes Natural Resources and Land Planning merge the concept models of GB/T 32853-2016, GB/T 40765-2021, and CityGML, yielding a unified taxonomy of terrain, water bodies, protected zones, and land-use parcels.

Spatio-temporal fundamentals. Spatial position and temporal extent are modeled in accordance with CityGML and the Time Ontology in OWL, enabling every urban object to carry precise geographic and temporal metadata.

Sensors and managed components. Urban management facilities such as lampposts and charging piles are subclassed using GB/T 30428.2-2013 and GB/T 36625.5-2019, while sensor entities adopt the modular SSN/SOSA ontology to capture observations, actuators, and deployment contexts.



**Figure 2.** SMOF core entity modules and primary classes. The five dashed boxes represent the core ontology modules, each containing corresponding top-level classes (11 in total).

Society and special topic data. Beyond cataloging natural resources and physical infrastructure, urban service must also manage its primary administrative subjects—natural persons and organizations. Accordingly, we introduce a Social and Thematic module, augmented with specialized datasets to accommodate remote sensing imagery, POI layers, panel statistics, and other domain specific resources.

Through this layered design, the SMOF delivers a coherent yet extensible schema capable of aligning heterogeneous standards, integrating redundant data models, and supporting a broad spectrum of smart city applications. Moreover, entity definitions can be further extended within specific application scenarios, enabling domain-oriented refinement without compromising the integrity of the universal hierarchy.

### 3.4. Attribute and Relation Design

To achieve a unified data vocabulary and promote inter-departmental collaboration, the ontology must specify not only what entities exist but also which characteristics they carry and how they interact [52]. In the SMOF, attributes capture intrinsic or extrinsic features of an entity, whereas relations embed entities into a coherent semantic network. Building on the modeling principles above, the SMOF adopts a dual-tier attribute–relation strategy that combines a universal semantic core with a thematic extension mechanism.

To balance generality with domain specificity, attributes are divided into a universal set and an extended set. The universal set forms the semantic core and is applicable to all classes, ensuring interoperability across domains. Extended attributes are introduced only when a particular application scenario requires additional detail. Following this layered strategy, three categories of universal attributes are defined: (i) basic, recording

administrative identity and provenance (e.g., GlobalId, DataSource, Name); (ii) temporal, capturing lifecycle aspects (e.g., UpdateDate, OriginalDate, NumericDuration); and (iii) spatial, describing geolocation and geometry (e.g., AbsoluteSpatialPosition, AreaCode, Dimension, SpatialSemanticDescription). Extended attributes supplement these categories when domain-specific precision is needed—for example, lane traffic volume in highway monitoring.

Relations in SMOF fall into eight macro-categories—Hierarchy, Part–Whole, Dependency, Functional, Spatial, Temporal, Sense, and Synonym—to maximize expressiveness while minimizing redundancy. Universal relations (e.g., isPartOf, adjacentTo, dependsOn) provide the backbone of inter-entity connectivity; domain-specific extensions (e.g., regulatedByPlan in planning, monitoredBySensor in emergency management) add finer granularity when required. In practice, universal relations can be aligned to domain-specific ones (e.g., hasIdentity  $\leftrightarrow$  hasCitizenID).

The same universal attribute or relation may carry different semantics, scopes, or value domains across contexts. For example, Status may denote typical congestion in transportation, whereas in building management it may indicate maintenance state. We therefore use a context-sensitive mechanism: define Status uniformly at the universal layer, then refine it into OperationalStatus, BuildingUsageStatus, and similar variants within specific domains, avoiding over-specification of the universal layer.

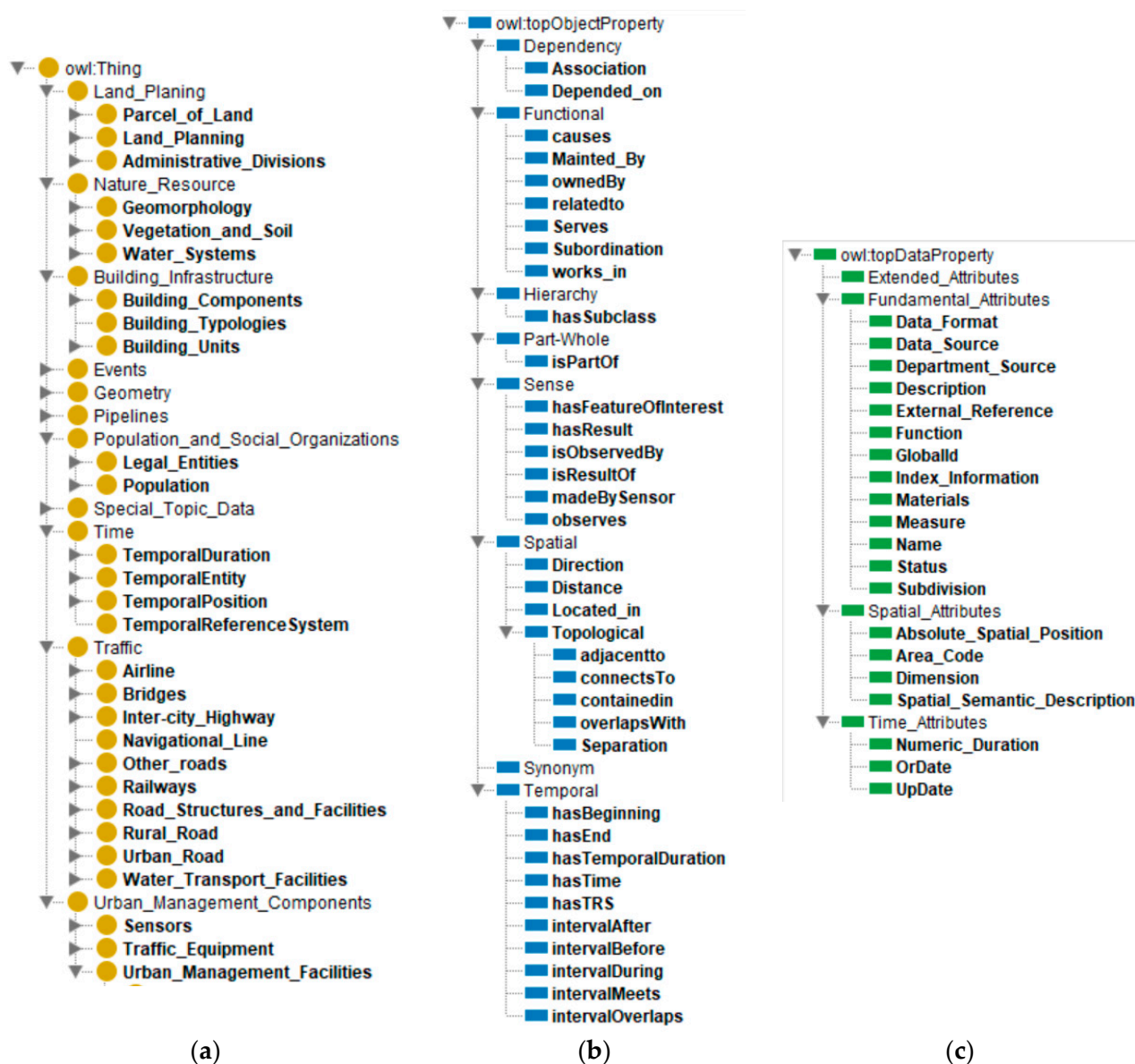
This dual-tier design yields a systematic model: universal components ensure cross-domain consistency, while extensions adapt to specific urban services. Combining a stable semantic core with scenario-driven expansion achieves conceptual generality and practical specificity, supporting reasoning and querying over heterogeneous urban data.

### 3.5. *Ontology Modeling and Mapping*

With the hierarchical classes, universal attributes, and universal relations in place, we instantiated the SMOF in Protégé. Inside the editor, Classes represent the urban-entity categories defined above, Object properties capture inter-entity relations, and Data properties contain the attributes. Figure 3 depicts the resulting ontology structure.

The urban ecosystem already hosts a multitude of ontologies and heterogeneous data sources. Aligning existing ontologies and data sources with SMOF and transforming data into SMOF instances constitute a key focus of current research. Consider the concept Sensor: it is defined in both ifcOWL and SOSA, but with different class names and storage conventions; likewise, IFC data are usually held in STEP files, CityGML in XML, and IoT feeds in JSON. Figure 4 sketches how typical urban ontologies map to the SMOF and how multi-source data are lifted into ontology instances:

(1) *Ontology-to-ontology alignment.* This refers to the process of establishing mappings between classes, attributes, and relations across different ontologies so that their semantics can interoperate consistently. Mapping rules can be written in the Semantic Web Rule Language (SWRL), which combines description logic with rule syntax. As illustrated in Figure 4, content from ifcOWL, SOSA, CityGML-RDF, and Time-OWL can all be translated into the SMOF. Following the literature [33], we distinguish three mapping types—direct, indirect, and attribute/relation, as summarized in Table 3.



**Figure 3.** SMOF modeling results in Protégé. (a) Entity modeling results (classes); (b) Relation modeling results (object properties); (c) Attribute modeling results (data properties).

(2) Data-to-ontology extraction (also referred to as knowledge lifting). Knowledge lifting denotes the transformation of heterogeneous raw data into ontology-compliant instances. On the data side, we usually use data extraction software or write code in Python 3.13.2 to achieve the parsing of heterogeneous data and convert it into instances of the ontology. For IFC, the `ifcopenshell` library identifies `IfcWall` objects, retrieves direct attributes such as `GlobalId`, and then follows links to `IfcPropertySet` to harvest non-geometric attributes (`IsExternal`, `Material`, `Thickness`). The collected facts are recorded as instances of `smof:Wall`. Analogous pipelines convert IoT JSON feeds, CityGML XML documents, or relational database tables into the SMOF individuals.

Through this dual strategy of SWRL alignment and automated data lifting, SMOF can reconcile diverse urban ontologies, ingest heterogeneous datasets, and furnish a coherent knowledge base for smart city analytics.





#### 4.2. Evaluation and Comparison

To systematically evaluate the coverage and expressiveness of SMOF, we complemented the competency question testing with a mixed evaluation protocol. This protocol combines objective indicators and subjective judgments. Specifically, objective evaluation was conducted using quantifiable indicators from OntoMetrics [55], while subjective evaluation relied on expert scoring and the LLM as judge paradigm [56]. For benchmarking, we compared SMOF against the three major ontologies introduced in the related work section: CIMO, UKG, and KM4City. For baseline ontologies that are not publicly available, we reconstructed their core schemas based on published literature to ensure fair comparison.

##### 4.2.1. Objective Evaluation

In the quantitative evaluation, we selected three structural indicators widely used in ontology engineering: Attribute Richness ( $AR$ ), Inheritance Richness ( $IR$ ), and Relationship Richness ( $RR$ ). These metrics, respectively, capture the average number of attributes per class, the average number of subclasses per class, and the proportion of non-inheritance relations relative to all relations. Their formal definitions are as follows:

$$AR = \frac{|ATT|}{|C|} \quad (4)$$

where  $|ATT|$  is the total number of attributes associated with classes and  $|C|$  is the total number of classes.

$$IR = \frac{\sum_{C_i \in C} |H^C(C_1, C_i)|}{|C|} \quad (5)$$

where  $|H|$  denotes inheritance relations.

$$RR = \frac{|P|}{|H| + |P|} \quad (6)$$

where  $|P|$  denotes non-inheritance relations (e.g., object properties, equivalence, or disjointness).

In practice,  $AR$  reflects the richness of attributes per entity, enabling more fine-grained semantic description;  $IR$  indicates the depth of class hierarchies, supporting precise categorization; and  $RR$  captures the diversity of relations, offering more expressive links for reasoning and integration. It should be noted, however, that higher values are not inherently better—these indicators need to be interpreted in relation to the intended scope and application scenarios of the ontology.

Table 4 reports the comparative results for SMOF and the three baseline ontologies.

**Table 4.** The evaluation results of the four ontologies in the indicators of  $AR$ ,  $IR$ , and  $RR$ .

Ontology	Attribute Richness	Inheritance Richness	Relationship Richness
KM4City	0.421	1.190	0.165
UrbanKG	8.283	0.583	0.754
CIMO	1.850	0.986	0.131
SMOF	9.688	0.981	0.127

From these results, SMOF exhibits the highest attribute richness, with an average of nearly ten attributes per class. This indicates that SMOF provides a highly detailed characterization of entities at the schema level, allowing for fine-grained instantiation. In terms of inheritance richness, SMOF achieves a value close to CIMO (0.986 vs. 0.981), reflecting a well-structured hierarchy that balances breadth and depth. However, its relationship richness is relatively low (0.127). This reflects SMOF's context-sensitive modeling strategy:

a small set of generic relations is defined and then specialized in domain contexts. This design limits the diversity of explicitly declared relation types.

By contrast, KM4City, as a long-standing and large-scale urban ontology, shows higher inheritance richness (1.190) but limited attribute and relationship richness due to its broad and stable schema design. CIMO, grounded in BIM and GIS data, demonstrates moderate values across the three indicators, reflecting its narrower scope. UrbanKG, constructed from POI and street-view data, records a relatively high relationship richness (0.754) owing to its dense web of non-inheritance links, but its inheritance structure remains shallow.

It is important to emphasize that these metrics evaluate only the formal richness and distributional characteristics of ontology design, rather than the correctness or adequacy of domain modeling. Consequently, while SMOF scores strongly in AR and IR, and adopts a principled approach to RR, the actual expressive power of the ontology still requires qualitative verification through complementary expert judgment and application-oriented scenario modeling.

#### 4.2.2. Subjective Evaluation

To complement the objective results, we conducted a subjective evaluation through both LLM as judge and expert scoring. The LLM-based scoring primarily measured conceptual completeness and coverage, while the expert evaluation targeted real-world scenario fitness and clarity of expression. This dual-layer design ensures that both formal conceptual quality and practical applicability are taken into account.

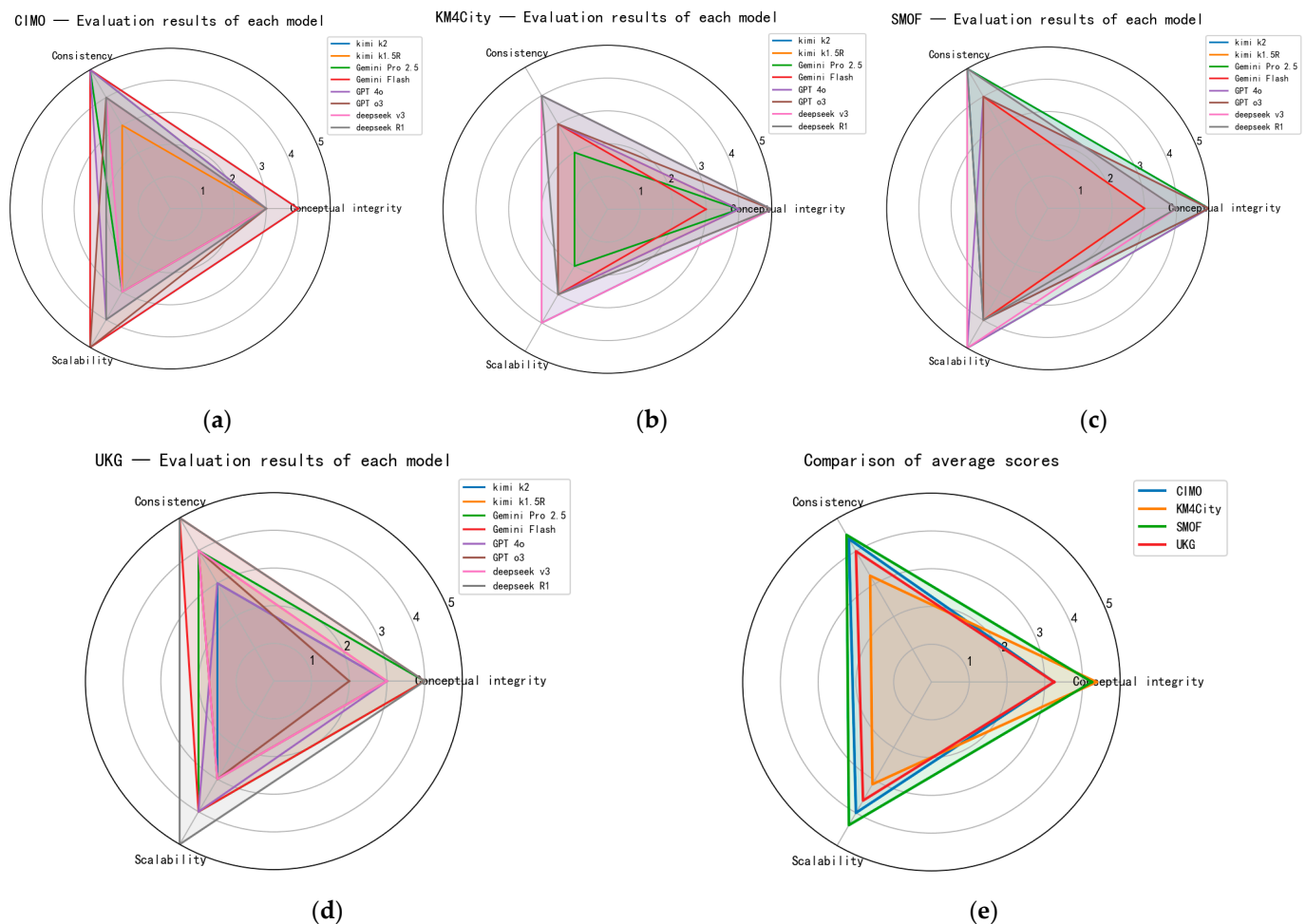
First, we employed the LLM as judge method to evaluate the overall conceptual coverage of SMOF. In this approach, LLMs were adopted because they combine broad coverage of general knowledge with substantial domain understanding [57]. Such models can detect whether a city ontology omits critical concepts, such as infrastructure types, transport networks, or public services. Unlike purely human scoring, they also apply consistent criteria and logical rules, which reduces subjectivity. Recently, LLM as Judge has also emerged as one of the most popular approaches for ontology evaluation, providing scalable and relatively objective assessment capabilities [58]. Specifically, we selected four representative LLMs (DeepSeek [59], GPT [60], Gemini [61] and Kimi [62]) as judges. The rationale for choosing these models is threefold. First, they jointly cover internationally recognized LLMs and those most widely adopted in our country, ensuring representativeness and authority. Second, they include both reasoning-oriented and text-oriented models, thus offering broader coverage. Third, all four models provide API access, which guarantees reproducibility and practical feasibility.

We evaluated three dimensions that reflect the structural and applicative characteristics of an ontology [63]. Concept completeness refers to the extent to which entities, attributes, and relations comprehensively cover the core concepts of the urban domain, thereby ensuring that the ontology does not omit essential knowledge. Concept consistency captures the freedom from internal contradictions in class definitions, property constraints, and relation logic, which is critical for maintaining logical validity and supporting reasoning tasks. Scalability denotes the ontology's capacity to accommodate new domains or entities without structural overhaul, reflecting its long-term adaptability and reusability in evolving smart city applications.

To clarify the evaluation methodology, we standardized the entire process at three stages. In the input stage, all ontologies were normalized into a JSON structure consisting of Entities, Attributes, and Relations. This ensured consistent input. In the prompt stage, the three evaluation dimensions were defined explicitly and rated on a five-point scale (1 = poor, 5 = excellent). The meaning of each score was described clearly in the prompt. In the output

stage, every task was repeated twice and the average taken, reducing randomness and inconsistency across models.

The evaluation results of LLM as judge are shown in Figure 5.



**Figure 5.** Radar charts of LLM-as-judge results and average scores. (a) Evaluation results for the CIMO model; (b) Evaluation results for the KM4City model; (c) Evaluation results for the SMOF model; (d) Evaluation results for the UKG model; (e) Comparison of average scores across all models. The three evaluation dimensions are defined as follows: completeness (coverage of domain concepts), consistency (absence of logical contradictions), and scalability (capacity to accommodate new domains or entities).

As summarized in Figure 5, completeness, consistency, and scalability jointly reflect domain coverage, logical validity, and adaptability. Averaging the three metrics across all models yields  $SMOF = 4.375$ , markedly higher than  $CIMO = 3.875$ ,  $UKG = 3.625$ , and  $KM4City = 3.583$ . We report the mean score as a balanced indicator of overall ontology quality, since completeness, consistency, and scalability are complementary dimensions. While individual scores reveal specific strengths and weaknesses, the average reflects the ontology's integrated conceptual soundness.

Although all four ontologies target the smart city domain, the score patterns reflect their differing design emphases. The SMOF excels in both completeness (4.25) and consistency (4.50); notably, Gemini Pro 2.5 awarded full marks on all three dimensions, indicating strong approval of the SMOF's layered entity taxonomy and its separation of universal versus extended attributes.

KM4City achieves the best completeness score (4.375) but fares worse in consistency (3.25) and scalability (3.125). Gemini Pro 2.5 rated it only 2 points for the latter two aspects, probably because KM4City's very large size and many near-duplicate concepts lack clear hierarchical documentation.

CIMO (3.875) and UKG (3.625) show balanced but lower performance. GPT-series models frequently assigned scores  $\leq 3$ , suggesting gaps in core entity definitions, attribute linkage, and relation coverage. This weakness is consistent with their original intentions: CIMO focuses on BIM/GIS/IoT interoperability within the CIM context, whereas UKG concentrates on street-view and POI analytics. Divergent problem statements thus translate into different strengths and weaknesses under the evaluation rubric.

The evaluation results of LLM as judge are shown in Figure 5. While this approach offers scalable and relatively objective assessments, it is not without limitations. The results may be influenced by model-specific biases and sensitivity to prompt phrasing, which can affect consistency across tasks. Therefore, we treated LLM-based scoring as complementary rather than definitive, and introduced an expert-based evaluation to provide qualitative insights and cross-validate the findings.

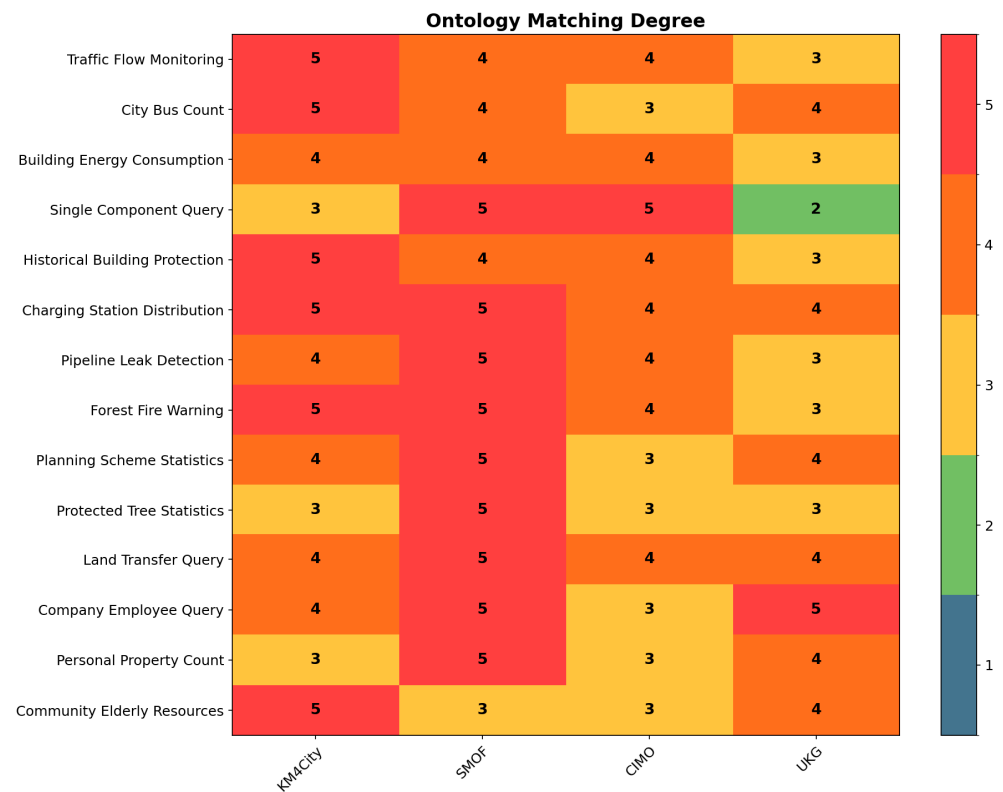
Subsequently we carried out an expert-based evaluation to complement the LLM study. Three specialists were invited: two experts in smart city applications and one in knowledge graph engineering. One expert first drafted fourteen coarse-grained application scenarios spanning five thematic domains: transport, buildings, urban infrastructure, urban planning, and society. Representativeness was ensured by referencing typical domains summarized in Wang et al.'s survey [34]. A second expert then attempted to model each scenario with SMOF, CIMO, KM4City and UKG, using only the native classes, properties and relations of each ontology and extending them only when absolutely necessary. Finally, a third expert assessed the resulting models on two criteria. The first was ontology–scenario fitness (ontology matching degree), meaning the extent to which the ontology provides the required vocabulary (entities, attributes, relations) without extension. The second was expression clarity, meaning the degree to which the chosen constructs convey the scenario unambiguously and consistently.

Both criteria were rated on a five-point scale (1 = poor, 5 = excellent). Coarse-grained scenarios were chosen intentionally so as not to penalize ontologies for lacking very fine detail.

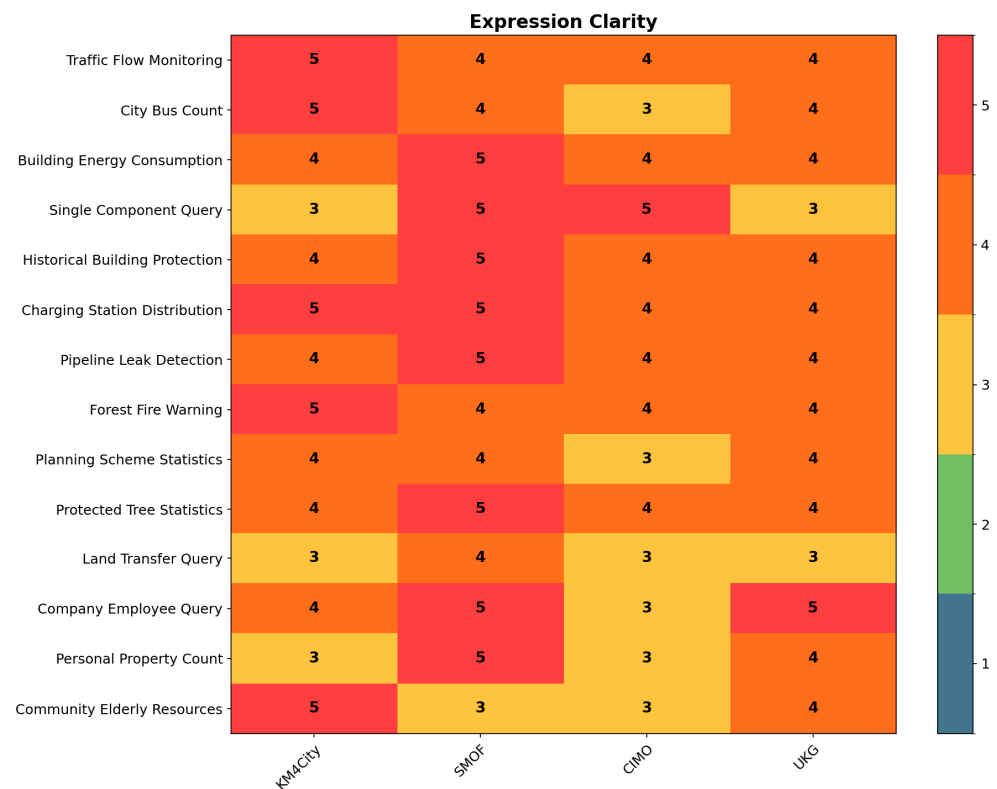
The heat map in Figure 6 summarizes the results. The SMOF achieved full marks for ontology scenario fitness in nine of the fourteen scenarios, including building component queries, charging station distribution, pipeline leak detection and forest fire early warning, demonstrating its strong alignment with a broad spectrum of urban tasks. KM4City ranked second, excelling in the two transport scenarios (traffic flow monitoring and bus ridership counting), reflecting its transport-centric design. CIMO, built around BIM/GIS/IoT interoperability, handled building and planning situations well but fell short in social resource and traffic contexts. UKG rarely scored the maximum, yet because it classifies a vast number of POIs, it could satisfy most scenarios by extending only one or two classes.

For expression clarity shown in Figure 7, the SMOF again led the field, receiving perfect scores in eight scenarios such as building energy accounting and protected tree statistics. Its lowest score (3/5) occurred in the “community services for the elderly” scenario. In such contexts, the existing universal attributes (e.g., ID, name, status) are insufficient unless extended, since domain-specific properties like “number of elderly beds” or “average resident age” only arise in these specialized settings. Given that the CIMO ontology had to be reconstructed from the original publication and is consequently incomplete, its clarity scores were predictably modest.





**Figure 6.** Ontology matching degree heatmap. The horizontal axis lists SMOF and baseline ontologies, while the vertical axis represents different practical scenarios. Higher scores indicate stronger fitness of an ontology to express knowledge in the corresponding scenario.



**Figure 7.** Expression clarity heatmap. The horizontal axis shows SMOF and baseline ontologies, and the vertical axis lists different scenarios. Higher scores denote clearer and more unambiguous expression of domain knowledge (clarity) in the corresponding scenario.

Overall, the SMOF outperformed the three baseline ontologies on both metrics across most scenarios, confirming that its cross-domain adaptability and semantic precision translate into practical modeling advantages. The experiment also highlights the differing strengths of the four ontologies: POI-oriented UKG and transport-rich KM4City are less suited to building-centric queries, illustrating the functional diversity that still characterizes urban ontological resources.

#### 4.3. Practical Scenario Application

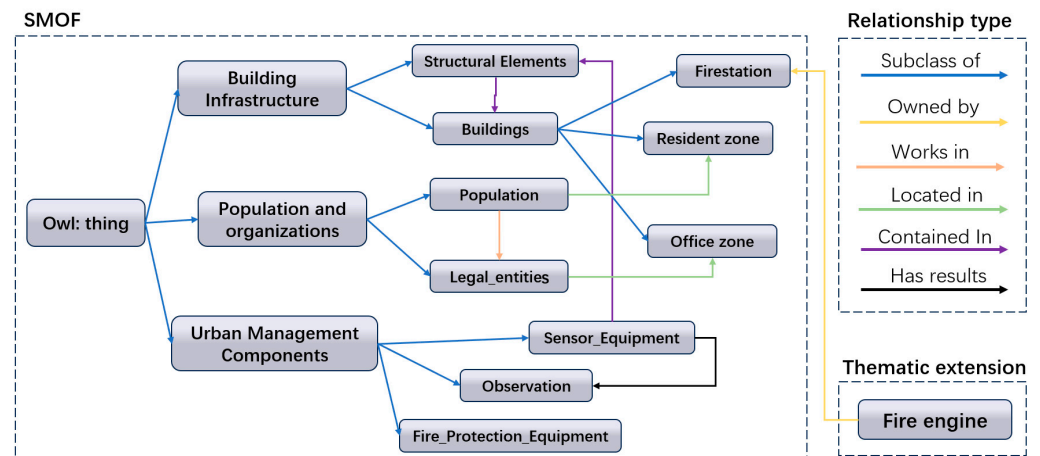
To ensure a comprehensive validation of SMOF, we employed two complementary urban service scenarios that highlight distinct aspects of ontology application. The fire-emergency case represents a high-stakes domain where timely warning and rapid decision-making rely on the seamless integration of heterogeneous real-time data. It demonstrates the ontology's capability for multi-source data representation and query execution. In contrast, the traffic-congestion case reflects a pervasive and socially impactful challenge in urban management, chosen to showcase the ontology's reasoning capacity for generating new knowledge beyond directly observed data. The datasets also differ in accessibility: the fire-emergency dataset, provided by Shenzhen Smart City Technology Development Co., Ltd. (Shenzhen, China) and the Shenzhen Spatial Digital Platform, cannot be publicly released due to contractual restrictions; meanwhile, the traffic-congestion dataset is available on figshare to support transparent conceptual validation.

Ontology evaluation must ultimately be grounded in urban service scenarios that stress heterogeneous data integration. Urban fire emergencies provide a representative testbed: timely warning and effective management demand rapid decisions based on sensor streams, building-information models, GIS layers, and relational data, yet linking observations to geospatial context and affected populations remains challenging [64]. Using SMOF as the semantic backbone, we carried out three steps: (i) formalized the knowledge required for fire response, (ii) lifted data from disparate sources into ontology instances, and (iii) executed graph queries over the resulting knowledge graph.

We extended SMOF with a fire-response module in Protégé (<http://protege.stanford.edu/>, accessed on 1 October 2025), adding classes such as FireEngine while reusing SMOF entities, attributes, and relations (Figure 8). Sensors are linked to their observations through `hasResult`. When a temperature reading exceeds a predefined threshold, the `ContainedIn` relation automatically triggers the association with the specific building component affected. From there, the model retrieves the building's address, locates nearby fire stations and hydrants, and notifies both office units and residents associated with the endangered location.

After modeling, we parsed IoT feeds, CityGML geometries, and relational tables and lifted them into SMOF ontology instances, persisting the results in a graph database. This knowledge lifting process, which extracts and semantically enriches multi-source heterogeneous data, forms a SMOF-based knowledge graph. The graph provides a unified, machine-readable representation of entities, their attributes (literals), and inter-entity relations. The knowledge graph supports user-defined SPARQL graph queries for relationship-centric retrieval and analysis.

To illustrate, we implemented four SPARQL patterns that meet core fire-response needs, as shown in Table 5. Temperature sensors with readings above 60 °C were selected as a demonstrative case to provide a clear numeric threshold. Through alignment with the SSN/SOSA framework, the same patterns can query other fire-relevant sensors (e.g., smoke or gas detectors) in real deployments. The patterns are executed directly against the triple store hosting the SMOF-based graph; the engine matches graph patterns over RDF triples and returns a binding table for entities satisfying the specified constraints.



**Figure 8.** Knowledge representation for fire emergency scenarios. Rectangles represent classes involved in knowledge modeling, and lines indicate relationships among entities. The types of relationships are described in the dashed box.

**Table 5.** SPARQL query in fire warning and response scenarios.

Objective	SPARQL Query
Find sensors whose temperature reading exceeds 60 °C	<pre> SELECT ?sensor WHERE {   ?sensor a smof:Sensor_Equipment;   smof:hasResult ?observation.   ?observation a smof:Observation;   smof:measure "°C";   smof:value ?value.   FILTER (?value &gt; 60) }</pre>
Identify the affected structural element and its parent building	<pre> SELECT ?structuralElement ?building WHERE {   ?sensor a smof:Sensor_Equipment.   ?sensor smof:ContainedIn ?structuralElement.   ?structuralElement a ?structSubClass.   ?structSubClass rdfs:subClassOf smof:StructuralElements.   ?structuralElement smof:isPartOf ?building.   ?building a ?bldgSubClass.   ?bldgSubClass rdfs:subClassOf smof:BuildingTypologies. }</pre>
Retrieve populations and companies located in the damaged building	<pre> SELECT ?populationName ?legalEntityName WHERE {   ?building a ?bldgSubClass.   ?bldgSubClass rdfs:subClassOf smof:BuildingTypologies.   ?population a smof:Population;   smof:works_in ?building;   smof:Name ?populationName.   ?legalEntity a smof:legalEntities;   smof:located_in ?building;   smof:Name ?legalEntityName. }</pre>

Table 5. Cont.

Objective	SPARQL Query
List fire protection equipment, fire stations and fire truck names within 1 km of the building	<pre> SELECT ?fireProtectionEquipment ?firestation ?fireTruckName WHERE {     ?building a ?bldgSubClass.     ?bldgSubClass rdfs:subClassOf smof:BuildingTypologies.     ?building smof:AbsoluteSpatialPosition ?buildingPos.     # equipment within 1 km     ?fireEquipment a smof:FireProtectionEquipment;     smof:AbsoluteSpatialPosition ?fireEqPos.     FILTER (geof:distance(?buildingPos, ?fireEqPos) &lt;= 1000)     # optional: fire station within 1 km     OPTIONAL {         ?firestation a smof:Firestation;         smof:AbsoluteSpatialPosition ?firestationPos.         FILTER (geof:distance(?buildingPos, ?firestationPos) &lt;= 1000)         ?firestation smof:ownedBy ?fireTruck.         ?fireTruck a smof:FireTruck;         smof:Name ?fireTruckName.     } }</pre>

We deployed the solution as a semantic web application on the Shenzhen Spatial Digital Platform using Java and the Apache Jena stack (Figure 9). The interface integrates BIM/GIS baselines with live sensor streams to support sensor search, real-time inspection, and impact-area analysis, powered by the ontology-backed knowledge graph.

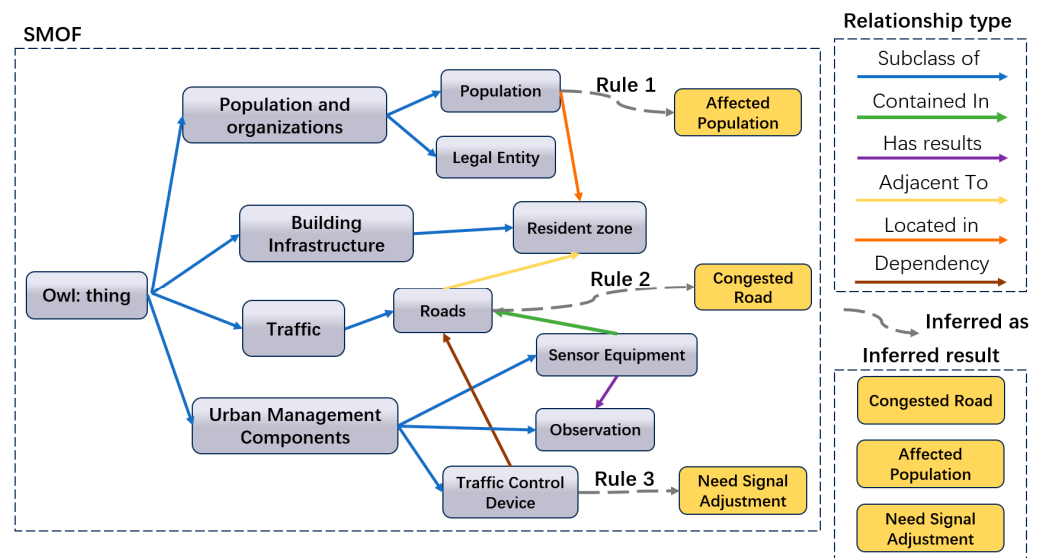


**Figure 9.** Fire warning and response interface display. Under “Overall interface display”: The left side is the emergency-related function panel, the middle shows scene visualization, and the right side displays specific functions.

Traffic congestion is a pervasive urban challenge with direct impacts on mobility, energy use, and public well-being [65]. To test transferability and reasoning capability, we include a compact, lightweight conceptual traffic example.

Within existing SMOF modules (no new extensions), we model Roads, Residential\_Zones, Populations, Sensor\_Equipment, Observations, and TrafficControlDevices

(Figure 10). On this basis, we define three inferred classes: CongestedRoad, AffectedPopulation, and NeedSignalAdjustment. These are derived by rules rather than asserted a priori.



**Figure 10.** Knowledge representation for traffic congestion scenarios. Rectangles represent classes involved in knowledge modeling, and lines indicate relationships among entities. The types of relationships are described in the dashed box.

SWRL extends OWL with Horn-like rules (antecedent  $\Rightarrow$  consequent) that combine class/property predicates with built-ins to infer new facts from asserted data [33,66]. Using SWRL we encode three rules that (i) classify congestion from speed/occupancy thresholds, (ii) link congested roads to affected populations via spatial containment, and (iii) flag devices requiring signal retiming (as shown in Table 6).

**Table 6.** SWRL Knowledge Reasoning in Traffic Congestion Scenarios.

Rule ID	Purpose	SWRL Rule	Results
1	Identify congested road segments from sensor observations	<pre> smof:Sensor_Equipment(?s) ^ smof:ContainedIn(?s, ?r) ^ smof:Road(?r) ^ smof:hasResult(?s, ?o1) ^ smof:Observation(?o1) ^ smof:measure(?o1, "vehicleCount"^^xsd:string) ^ smof:value(?o1, ?c) ^ swrlb:greaterThan(?c, 100) ^ smof:hasResult(?s, ?o2) ^ smof:Observation(?o2) ^ smof:measure(?o2, "avgSpeed"^^xsd:string) ^ smof:value(?o2, ?v) ^ swrlb:lessThan(?v, 20) -&gt; smof:Congested_Road(?r) </pre>	smof:Congested_Road



Table 6. Cont.

Rule ID	Purpose	SWRL Rule	Results
2	Propagate road congestion to human impact via spatial context	$\text{smof:Congested\_Road}(?r) \wedge$ $\text{smof:adjacentTo}(?r, ?z) \wedge$ $\text{smof:Resident\_zone}(?z) \wedge$ $\text{smof:Population}(?p) \wedge$ $\text{smof:located\_in}(?p, ?z)$ $\rightarrow \text{smof:Affected\_Population}(?p)$	smof:Affected_Population
3	Recommend operational adjustment for dependent devices	$\text{smof:Congested\_Road}(?r) \wedge$ $\text{smof:Traffic\_Control\_Device}(?d) \wedge$ $\text{smof:Dependency}(?d, ?r)$ $\rightarrow \text{smof:Need\_Signal\_Adjustment}(?d)$	smof:Need_Signal_Adjustment

The modeling and reasoning process was implemented in Protégé. Classes, object properties, and data properties were defined according to the SMOF structure, and individuals were created to represent roads, sensors, observations, resident zones, and devices. The SWRL rules were encoded in the SWRLTab, and reasoning was performed with the Pellet reasoner. The system successfully inferred congested roads, identified affected populations, and flagged relevant control devices for adjustment. As summarized in Figure 11, the results confirm that SMOF can support not only semantic retrieval of heterogeneous urban data but also knowledge driven inference that enriches decision making.

**Protégé SWRL reasoning Tab**

**Inferred Class Hierarchy**

**Inferred Types of Individual r\_main**

Figure 11. The reasoning process and results in the Protégé.

Collectively, the results confirm that semantic modeling atop SMOF enables both cross-source querying of heterogeneous urban data and prescriptive reasoning. This bridges data integration and intelligent urban governance.

## 5. Discussion

This study was guided by three questions—RQ1 (how to achieve semantic integration of multi-source heterogeneous urban data), RQ2 (how to construct an ontology that systematically encompasses cross-domain urban management objects), and RQ3 (how to validate such a framework in real scenarios)—together with three corresponding hypotheses (H1–H3). The findings confirm all three hypotheses. For RQ1/H1, alignment with authoritative standards, competency questions, and explicit modeling principles jointly reduced semantic fragmentation and representational inconsistency. Authoritative standards provided canonical definitions of classes, attributes, and relations; the modeling principles regulate scope, granularity, naming, and classification so that outcomes remain faithful to requirements while interoperable with external sources; and CQs translated requirements into testable queries that the ontology must satisfy after construction, ensuring that modeling outcomes remained consistent with the defined requirements. This mechanism is also reflected in the metrics: SMOF showed the highest AR and consistency scores and balanced IR/RR values (Figures 5–7), evidencing reduced semantic gaps and more coherent representation than the baselines. For RQ2/H2, the combination of a layered taxonomy and a dual-tier attribute–relation design enabled cross-domain modeling. Universal attributes and relations formed a semantic core, capturing the ‘greatest common denominator’ across diverse urban contexts. Extended attributes and relations preserved flexibility, enabling domain-specific details without compromising the core. Context-sensitive specialization further ensured that universal elements remained reusable across various domains. For RQ3/H3, effectiveness and applicability were established through a mixed evaluation protocol—combining structural indicators, LLM as judges, and expert scoring—together with two empirical applications (fire emergency and traffic reasoning). These jointly tested formal soundness, conceptual coverage, scenario fitness, and operational practicality.

The comparative experiments highlight SMOF’s advantages in both structure and applicability. In the quantitative evaluation, SMOF achieved the highest attribute richness (9.688), indicating a detailed schema with rich descriptive capacity, and maintained an inheritance richness of 0.98, comparable to CIMO’s 0.99 but with broader coverage. In the LLM as judge evaluation, SMOF obtained mean scores of 4.25 for completeness, 4.50 for consistency, and 4.38 for scalability, surpassing KM4City, UrbanKG, and CIMO. Expert assessments corroborated these findings, showing that SMOF achieved scenario fitness scores of 5/5 in nine out of fourteen representative applications, while other ontologies required more extensions. Moreover, two empirical studies reinforced the framework’s utility: the fire emergency case demonstrated its ability to integrate sensors, buildings, and stations for real-time heterogeneous data query support, while the traffic congestion case illustrated its reasoning capacity to infer congestion states, affected populations, and signal adjustments.

Ontology evaluation, however, remains a recognized challenge. No universally accepted framework currently exists for measuring expressiveness, and each method carries limitations. Quantitative indicators are objective but insufficient for capturing knowledge representation capacity. LLM-based evaluation provides scalability and domain breadth, but its outputs can be unstable. Expert judgment is grounded in domain knowledge yet is inevitably subjective and resource-intensive. Nevertheless, this study combines four approaches: quantitative metrics, LLM-based evaluation, expert scoring, and empirical validation, thereby providing a comprehensive evaluation protocol. This multi-layered strategy strengthens the credibility of the findings despite inherent limitations. Future studies will extend SMOF to additional urban service scenarios to further verify its knowledge representation capacity, and will involve a larger, more diverse group of experts with complementary evaluation methods to ensure more accurate assessments.

Relative to our 2024 preprint version [11], this work represents substantial progress. The SMOF has evolved from a preliminary conceptual design into a refined, layered framework with extensive semantic coverage. Evaluation has expanded from single scenario validation to a multi-dimensional protocol integrating quantitative, qualitative, and empirical perspectives. The ontology itself has been enriched with universal and extended attributes, refined relations, and broader mappings, thereby enhancing its expressiveness and extensibility. These advances transform SMOF from an initial idea into a robust and mature framework for smart city knowledge integration.

## 6. Conclusions

This study proposed the SMOF, a hierarchical, standards-driven ontology designed to unify heterogeneous urban data. The framework integrates five core entity modules, eleven major entity categories, and a dual-tier relation scheme, ensuring both conceptual generality and domain adaptability. Comparative evaluations indicated that SMOF exhibits strong structural soundness, rich conceptual representation, and robust knowledge expressiveness compared with baseline ontologies. Empirical validations in fire emergency management and traffic congestion further confirmed its ability to integrate multi-source data, enable semantic reasoning, and generate actionable insights for urban governance. These results collectively confirm the three hypotheses and provide positive answers to the guiding research questions.

Despite these contributions, several limitations remain. First, the empirical validation is restricted to two scenarios—fire emergency and traffic congestion—which, although representative, do not fully reflect the diversity of smart city services such as land-use planning, environmental monitoring, or social management. Second, ontology construction for specific scenarios is still labor-intensive and requires extensive expert involvement, which hinders rapid deployment across new domains. Third, the multi-faceted evaluation protocol, while comprehensive, also has shortcomings: quantitative metrics mainly capture structural properties, LLM as judge results may be affected by model bias and prompt sensitivity, and expert-based scoring is limited by the small sample size.

These limitations directly inform our future research agenda. To reduce manual effort, we plan to integrate LLMs with retrieval-augmented generation (RAG) to support semi-automated ontology construction and scenario adaptation [67,68]. To address the static-data constraint, we will extend SMOF to incorporate dynamic data streams, enabling real-time reasoning and anomaly detection (e.g., sudden traffic or infrastructure changes). To broaden applicability, we aim to validate SMOF in additional domains such as natural resource management and land-use planning. Finally, to enhance the robustness of evaluation, we will expand expert participation and explore hybrid validation methods that combine human judgment with automated assessment.

## 7. Patents

An invention patent application derived from this work has been submitted: Application Number 2025108046798, filed on 17 June 2025, entitled “A Method, System, Terminal, and Storage Medium for Constructing a Knowledge Graph of Heterogeneous Data from Multiple Urban Departments”.

**Author Contributions:** Conceptualization, X.H.; Methodology, X.K. and R.G.; Software, B.H.; Validation, X.K.; Formal analysis, X.L.; Resources, Z.Q.; Writing—original draft preparation, X.H.; Writing—review and editing, X.H.; Visualization, B.H.; Supervision, R.G.; Funding acquisition, X.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China, grant number 41901325; and the National Key Research and Development Program of China, grant number 2022YFC3800604.

**Data Availability Statement:** The datasets generated during and analyzed during the current study are publicly available at: <https://figshare.com/s/2875e770d9907435b1fb>, accessed on 1 October 2025.

**Acknowledgments:** During the preparation of this study, the authors used GPT, Deepseek, Kimi and Gemini for the purposes of LLM as judge evaluation. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Appendix A

Table A1 presents the detailed competency questions and corresponding Owl format answers used to validate SMOF's coverage, consistency, and interoperability.

**Table A1.** Competency questions and illustrative OWL fragments.

Competency Question	Owl Format Answers
CQ1	# Hierarchy (subClassOf) smof:Building_components rdfs:subClassOf smof:Building_infrastructure. # Intrinsic information (data property) smof:hasName a owl:DatatypeProperty; rdfs:domain:Entity; rdfs:range xsd:string. # Inter-entity relationship (object property) smof:adjacentTo a owl:ObjectProperty; rdfs:domain smof:SpatialEntity; rdfs:range smof:SpatialEntity.
	# Top-level classes:SMOF smof:SMOF a owl:Class. smof:Building_infrastructure a owl:Class; rdfs:subClassOf smof:SMOF. smof:Events a owl:Class; rdfs:subClassOf smof:SMOF. smof:Geometry a owl:Class; rdfs:subClassOf smof:SMOF. smof:Nature_and_Geographic_Space a owl:Class; rdfs:subClassOf smof:SMOF. smof:Pipelines a owl:Class; rdfs:subClassOf smof:SMOF. smof:Population_and_Social_Organizations a owl:Class; rdfs:subClassOf smof:SMOF. smof:Special_Topic_Data a owl:Class; rdfs:subClassOf smof:SMOF. smof:Time a owl:Class; rdfs:subClassOf smof:SMOF. smof:Traffic a owl:Class; rdfs:subClassOf smof:SMOF. smof:Urban_Management_Components a owl:Class; rdfs:subClassOf smof:SMOF.
CQ2	

Table A1. Cont.

Competency Question	Owl Format Answers
CQ3	<pre> # Spatial-attribute super-class smof:spatial_attributes a owl:Class. # Address (semantic description) smof:Spatial_semantic_description a   owl:DatatypeProperty; rdfs:subPropertyOf smof:spatial_attributes; rdfs:domain smof:SpatialEntity; rdfs:range xsd:string. # Absolute position (lat-lon) smof:AbsoluteSpatialPosition a owl:DatatypeProperty; rdfs:subPropertyOf smof:spatial_attributes; rdfs:domain smof:SpatialEntity; rdfs:range xsd:string. # Topological relations spatial:topological a owl:Class. smof:adjacentTo a owl:ObjectProperty; rdfs:subPropertyOf spatial:topological; rdfs:domain smof:SpatialEntity; rdfs:range smof:SpatialEntity. smof:connectsTo a owl:ObjectProperty; rdfs:subPropertyOf spatial:topological; rdfs:domain SpatialEntity; rdfs:range SpatialEntity. smof:ContainedIn a owl:ObjectProperty; rdfs:subPropertyOf spatial:topological; rdfs:domain SpatialEntity; rdfs:range SpatialEntity. smof:overlapsWith a owl:ObjectProperty; rdfs:subPropertyOf spatial:topological; rdfs:domain SpatialEntity; rdfs:range SpatialEntity. smof:Separation a owl:ObjectProperty; rdfs:subPropertyOf spatial:topological; rdfs:domain smof:SpatialEntity; rdfs:range smof:SpatialEntity. </pre>
CQ4	<pre> # Alignment with KM4City smof:Administration a owl:Class; owl:equivalentClass km4c:Administration. smof:LocalPublicTransport a owl:Class; owl:equivalentClass km4c:Localpublictransport. smof:Sensors a owl:Class; owl:equivalentClass km4c:Sensors. # Alignment with UrbanKG smof:POIs a owl:Class; owl:equivalentClass ukg:POIs. smof:Users a owl:Class; owl:equivalentClass ukg:Users. smof:Satellite_Images a owl:Class; owl:equivalentClass ukg:Satellite images. smof:Street_View_Images a owl:Class; owl:equivalentClass ukg:Street view images. </pre>



Table A1. Cont.

Competency Question	Owl Format Answers
CQ5	# Alignment with CityGML smof:Geometry a owl:Class; owl:equivalentClass citygml:Geometry. smof:Building_Geometry a owl:Class; owl:equivalentClass citygml:Building.
CQ6	# Alignment with SOSA smof:Sensor a owl:Class; owl:equivalentClass sosa:Sensor. smof:Observes a owl:ObjectProperty; owl:equivalentProperty sosa:observes.
CQ7	# Dependency relations smof:Dependency a owl:Class. smof:Association a owl:ObjectProperty; rdfs:subPropertyOf smof:Dependency; rdfs:domain smof:Entity; rdfs:range smof:Entity. smof:Depended_on a owl:ObjectProperty; rdfs:subPropertyOf smof:Dependency; rdfs:domain smof:Entity; rdfs:range smof:Entity. # Whole-part relations smof:Part-Whole a owl:Class. smof:isPartOf a owl:ObjectProperty; rdfs:subPropertyOf smof:Part-Whole; rdfs:domain smof:Entity; rdfs:range smof:Entity.
CQ8	# Transitive whole-part reasoning smof:isPartOf a owl:TransitiveProperty. # Transitive containment:containedIn smof:ContainedIn a owl:TransitiveProperty. # Class-hierarchy reasoning smof:Residential_Building a owl:Class; rdfs:subClassOf smof:Building_infrastructure.
CQ9	# Fundamental attribute set smof:Fundamental_Attributes a owl:Class. # Typical data properties smof:Data_Source a owl:DatatypeProperty; rdfs:subPropertyOf smof:Fundamental_Attributes; rdfs:domain smof:Entity; rdfs:range xsd:string. smof:Description a owl:DatatypeProperty; rdfs:subPropertyOf smof:Fundamental_Attributes; rdfs:domain smof:Entity; rdfs:range xsd:string. smof:Name a owl:DatatypeProperty; rdfs:subPropertyOf smof:Fundamental_Attributes; rdfs:domain smof:Entity; rdfs:range xsd:string. smof:Status a owl:DatatypeProperty; rdfs:subPropertyOf smof:Fundamental_Attributes; rdfs:domain smof:Entity; rdfs:range xsd:string.

Table A1. Cont.

Competency Question	Owl Format Answers
CQ10	# Globally unique identifier smof:hasGlobalID a owl:DatatypeProperty; rdfs:domain smof:Entity; rdfs:range xsd:string; owl:functionalProperty "true"^^xsd:boolean.

## References

1. Urbanization. Available online: <https://www.unfpa.org/urbanization#readmore-expand> (accessed on 1 October 2025).
2. Alahi, M.E.E.; Sukkuea, A.; Tina, F.W.; Nag, A.; Kurdthongmee, W.; Suwannarat, K.; Mukhopadhyay, S.C. Integration of IoT-enabled technologies and artificial intelligence (AI) for smart city scenario: Recent advancements and future trends. *Sensors* **2023**, *23*, 5206. [\[CrossRef\]](#)
3. Panagiotopoulou, M.; Stratigea, A.; Kokla, M. Smart, Sustainable, Resilient, and Inclusive Cities: Integrating Performance Assessment Indicators into an Ontology-Oriented Scheme in Support of the Urban Planning Practice. *Urban Sci.* **2025**, *9*, 33. [\[CrossRef\]](#)
4. Cheshmehzangi, A.; Batty, M.; Allam, Z.; Jones, D.S. *City Information Modelling*; Springer: Berlin/Heidelberg, Germany, 2024.
5. Han, M.J.N.; Kim, M.J. A systematic review of smart city research from an urban context perspective. *Cities* **2024**, *150*, 105027. [\[CrossRef\]](#)
6. Del Campo, G.; Saavedra, E.; Piovano, L.; Luque, F.; Santamaria, A. Virtual reality and internet of things based digital twin for smart city cross-domain interoperability. *Appl. Sci.* **2024**, *14*, 2747. [\[CrossRef\]](#)
7. Pliatsios, A.; Kotis, K.; Goumopoulos, C. A systematic review on semantic interoperability in the IoE-enabled smart cities. *Internet Things* **2023**, *22*, 100754. [\[CrossRef\]](#)
8. Dospinescu, O.; Perca, M. Technological integration for increasing the contextual level of information. *Analele Stiintifice Ale Univ. Alexandru Ioan Cuza Din Iasi-Stiinte Econ.* **2011**, *58*, 571–581.
9. Li, T.; Rui, Y.; Zhu, H.; Lu, L.; Li, X. Comprehensive digital twin for infrastructure: A novel ontology and graph-based modelling paradigm. *Adv. Eng. Inform.* **2024**, *62*, 102747. [\[CrossRef\]](#)
10. Hogan, A.; Blomqvist, E.; Cochez, M.; d’Amato, C.; Melo, G.D.; Gutierrez, C.; Kirrane, S.; Gayo, J.E.L.; Navigli, R.; Neumaier, S. Knowledge graphs. *ACM Comput. Surv.* **2021**, *54*, 1–37. [\[CrossRef\]](#)
11. Kuai, X.; He, X.; He, B.; Liu, Y.; Zhao, Z.; Guo, R. Smart City Ontology Framework for Urban Data Integration and Governance Applications. *Preprints* **2024**.
12. Hao, X.; Chen, W.; Yan, Y.; Zhong, S.; Wang, K.; Wen, Q.; Liang, Y. Urbanvlp: Multi-granularity vision-language pretraining for urban socioeconomic indicator prediction. In Proceedings of the AAAI Conference on Artificial Intelligence, Philadelphia, PA, USA, 25 February–4 March 2025; pp. 28061–28069.
13. Deng, Y.; Zhao, X.; Sun, H.; Chen, Y.; Wang, X.; Xue, X.; Li, L.; Song, J.; Hsieh, C.-Y.; Hou, T. RSGPT: A generative transformer model for retrosynthesis planning pre-trained on ten billion datapoints. *Nat. Commun.* **2025**, *16*, 7012. [\[CrossRef\]](#)
14. Adegun, A.A.; Fonou-Dombeu, J.V.; Viriri, S.; Odindi, J. Ontology-Based Deep Learning Model for Object Detection and Image Classification in Smart City Concepts. *Smart Cities* **2024**, *7*, 2182–2207. [\[CrossRef\]](#)
15. Zou, X.; Yan, Y.; Hao, X.; Hu, Y.; Wen, H.; Liu, E.; Zhang, J.; Li, Y.; Li, T.; Zheng, Y.; et al. Deep learning for cross-domain data fusion in urban computing: Taxonomy, advances, and outlook. *Inf. Fusion* **2025**, *113*, 102606. [\[CrossRef\]](#)
16. Choi, S.; Yoon, S. AI agent-based intelligent urban digital twin (I-UDT): Concept, methodology, and case studies. *Smart Cities* **2025**, *8*, 28. [\[CrossRef\]](#)
17. ŞAHİN, E.; Arslan, N.N.; Özdemir, D. Unlocking the black box: An in-depth review on interpretability, explainability, and reliability in deep learning. *Neural Comput. Appl.* **2025**, *37*, 859–965. [\[CrossRef\]](#)
18. Du, H.; Wei, L.; Dimitrova, V.; Magee, D.; Clarke, B.; Collins, R.; Entwisle, D.; Torbaghan, M.E.; Curioni, G.; Stirling, R. City infrastructure ontologies. *Comput. Environ. Urban Syst.* **2023**, *104*, 101991. [\[CrossRef\]](#)
19. Hofmeister, M.; Brownbridge, G.; Hillman, M.; Mosbach, S.; Akroyd, J.; Lee, K.F.; Kraft, M. Cross-domain flood risk assessment for smart cities using dynamic knowledge graphs. *Sustain. Cities Soc.* **2024**, *101*, 105113. [\[CrossRef\]](#)
20. Zeng, Z.; Qin, J.; Wu, T. A Knowledge Graph-Enhanced Hidden Markov Model for Personalized Travel Routing: Integrating Spatial and Semantic Data in Urban Environments. *Smart Cities* **2025**, *8*, 75. [\[CrossRef\]](#)
21. Liu, J.; Guo, D.; Liu, G.; Zhao, Y.; Yang, W.; Tang, L. Construction Method of City-Level Geographic Knowledge Graph Based on Geographic Entity. In Proceedings of the International Conference on Geoinformatics and Data Analysis, Lyon, France, 21–23 January 2022; pp. 133–142.

22. Huang, C.-Y.; Chiang, Y.-H.; Tsai, F. An ontology integrating the open standards of city models and Internet of things for smart-city applications. *IEEE Internet Things J.* **2022**, *9*, 20444–20457. [\[CrossRef\]](#)
23. Ding, L.; Xiao, G.; Pano, A.; Fumagalli, M.; Chen, D.; Feng, Y.; Calvanese, D.; Fan, H.; Meng, L. Integrating 3D city data through knowledge graphs. *Geo-Spat. Inf. Sci.* **2025**, *28*, 780–799. [\[CrossRef\]](#)
24. Grisiute, A.; Raubal, M.; Herthogs, P. 3D Land Use Planning: Making Future Cities Measurable with Ontology-Driven Representations of Planning Regulations. *Agil. GIScience Ser.* **2025**, *6*, 3. [\[CrossRef\]](#)
25. Tok, Y.C.; Zheng, D.Y.; Chattopadhyay, S. A Smart City Infrastructure ontology for threats, cybercrime, and digital forensic investigation. *Forensic Sci. Int. Digit. Investig.* **2025**, *52*, 301883. [\[CrossRef\]](#)
26. Abid, T.; Zarzour, H.; Laouar, M.R.; Khadir, M.T. Towards a smart city ontology. In Proceedings of the 2016 IEEE/ACS 13th International Conference of Computer Systems and Applications (AICCSA), Agadir, Morocco, 29 November–2 December 2016; pp. 1–6.
27. Joshi, S.; Saxena, S.; Godbole, T. Developing smart cities: An integrated framework. *Procedia Comput. Sci.* **2016**, *93*, 902–909. [\[CrossRef\]](#)
28. Qamar, T.; Bawany, N.Z.; Javed, S.; Amber, S. Smart city services ontology (SCSO): Semantic modeling of smart city applications. In Proceedings of the 2019 Seventh International Conference on Digital Information Processing and Communications (ICDIPC), Trabzon, Turkey, 2–4 May 2019; pp. 52–56.
29. De Nicola, A.; Villani, M.L. Smart city ontologies and their applications: A systematic literature review. *Sustainability* **2021**, *13*, 5578. [\[CrossRef\]](#)
30. Espinoza-Arias, P.; Poveda-Villalón, M.; García-Castro, R.; Corcho, O. Ontological representation of smart city data: From devices to cities. *Appl. Sci.* **2018**, *9*, 32. [\[CrossRef\]](#)
31. Pereira, A.P.; Prokopiuk, M. Informational Modeling of Cities: Method and Challenges for Institutional Implementation. *J. Urban Technol.* **2024**, *31*, 97–115. [\[CrossRef\]](#)
32. Xu, Z.; Guan, W.; Zhang, G.; Fang, Z.; Cai, W. Construction of the CIM hierarchical classification semantic network based on multi-source heterogeneous data. *J. Tsinghua Univ. (Sci. Technol.)* **2025**, *65*, 1197–1208. [\[CrossRef\]](#)
33. Shi, J.; Pan, Z.; Jiang, L.; Zhai, X. An ontology-based methodology to establish city information model of digital twin city by merging BIM, GIS and IoT. *Adv. Eng. Inform.* **2023**, *57*, 102114. [\[CrossRef\]](#)
34. Wang, Z.; Han, F.; Zhao, S. A Survey on Knowledge Graph Related Research in Smart City Domain. *ACM Trans. Knowl. Discov. Data* **2024**, *18*, 1–31. [\[CrossRef\]](#)
35. Liu, Y.; Ding, J.; Fu, Y.; Li, Y. Urbankg: An urban knowledge graph system. *ACM Trans. Intell. Syst. Technol.* **2023**, *14*, 1–25. [\[CrossRef\]](#)
36. Bellini, P.; Benigni, M.; Billero, R.; Nesi, P.; Rauch, N. Km4City ontology building vs data harvesting and cleaning for smart-city services. *J. Vis. Lang. Comput.* **2014**, *25*, 827–839. [\[CrossRef\]](#)
37. Fernández-López, M.; Gómez-Pérez, A.; Juristo Juzgado, N. Methontology: From ontological art towards ontological engineering. In Proceedings of the AAAI-97 Spring Symposium Series, Stanford, CA, USA, 24–26 March 1997.
38. Keet, C.M.; Khan, Z.C. Discerning and Characterising Types of Competency Questions for Ontologies. *arXiv* **2024**, arXiv:2412.13688. [\[CrossRef\]](#)
39. Ye, P.; Zhang, X.; Shi, G.; Chen, S.; Huang, Z.; Tang, W. TKRM: A formal knowledge representation method for typhoon events. *Sustainability* **2020**, *12*, 2030. [\[CrossRef\]](#)
40. GB/T 40765-2021; Ontology Model for Fundamental Geographic Information. State Administration for Market Regulation, Standardization Administration of China: Beijing, China, 2021.
41. GB/T 13923-2022; Classification and Codes for Fundamental Geographic Information Feature. State Administration for Market Regulation, Standardization Administration of China: Beijing, China, 2022.
42. Gröger, G.; Plümer, L. CityGML—Interoperable semantic 3D city models. *ISPRS J. Photogramm. Remote Sens.* **2012**, *71*, 12–33. [\[CrossRef\]](#)
43. CJJ/T157-2010; Technical Code for Three-Dimensional City Modelling. Ministry of Housing and Urban-Rural Development of the People's Republic of China: Beijing, China, 2018.
44. GB/T 30428.2-2013; Information System for Digitized Supervision and Management of City—Part 2: Managed Component and Event. General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of China: Beijing, China, 2013.
45. GB/T 28590-2012; Classification and Code for Urban Underground Facilities. General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of China: Beijing, China, 2012.
46. GB/T 36625.5-2019; Smart City—Data Fusion—Part 5: Data Elements of Basic Municipal Facilities. State Administration for Market Regulation, Standardization Administration of China: Beijing, China, 2019.
47. Pauwels, P.; Terkaj, W. EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Autom. Constr.* **2016**, *63*, 100–133. [\[CrossRef\]](#)

48. GB/T 51269-2017; Standard for Classification and Coding of Building Information Modeling. Ministry of Housing and Urban-Rural Development of the People's Republic of China: Beijing, China, 2017.
49. Janowicz, K.; Haller, A.; Cox, S.J.; Le Phuoc, D.; Lefrançois, M. SOSA: A lightweight ontology for sensors, observations, samples, and actuators. *J. Web Semant.* **2019**, *56*, 1–10. [\[CrossRef\]](#)
50. Pan, F.; Hobbs, J.R. Time ontology in owl. *W3C Work. Draft. W3C* **2006**, *1*, 1.
51. ISO 19126:2021; Geographic Information—Feature Concept Dictionaries and Registers. State Administration for Market Regulation, Standardization Administration of China: Beijing, China, 2016.
52. Noy, N.F. *Ontology Development 101: A Guide to Creating Your First Ontology*; Stanford Knowledge Systems Laboratory Technical Report, KSL-01-05; Stanford University: Stanford, CA, USA, 2001.
53. Ding, Y.; Xu, Z.; Zhu, Q.; Li, H.; Luo, Y.; Bao, Y.; Tang, L.; Zeng, S. Integrated data-model-knowledge representation for natural resource entities. *Int. J. Digit. Earth* **2022**, *15*, 653–678. [\[CrossRef\]](#)
54. Qing, Z.; Hankan, L.; Haowei, Z.; Mingwei, L.; Yulin, D.; Xiaochun, R.; Wei, W.; Liguang, Z.; Xun, L.; Jun, Z. Classification and Coding of Entity Features for Digital Twin Sichuan-Tibet Railway. *Geomat. Inf. Sci. Wuhan Univ.* **2020**, *45*, 1319–1327.
55. Lantow, B. Ontometrics: Putting metrics into use for ontology evaluation. In Proceedings of the KEOD, Porto, Portugal, 9–11 November 2016; pp. 186–191.
56. Kommineni, V.K.; König-Ries, B.; Samuel, S. From human experts to machines: An LLM supported approach to ontology and knowledge graph construction. *arXiv* **2024**, arXiv:2403.08345. [\[CrossRef\]](#)
57. Gu, J.; Jiang, X.; Shi, Z.; Tan, H.; Zhai, X.; Xu, C.; Li, W.; Shen, Y.; Ma, S.; Liu, H. A survey on llm-as-a-judge. *arXiv* **2024**, arXiv:2411.15594. [\[CrossRef\]](#)
58. Lippolis, A.S.; Saeedizade, M.J.; Keskiärrkkä, R.; Gangemi, A.; Blomqvist, E.; Nuzzolese, A.G. Large Language Models Assisting Ontology Evaluation. *arXiv* **2025**, arXiv:2507.14552. [\[CrossRef\]](#)
59. Guo, D.; Yang, D.; Zhang, H.; Song, J.; Zhang, R.; Xu, R.; Zhu, Q.; Ma, S.; Wang, P.; Bi, X. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv* **2025**, arXiv:2501.12948.
60. Achiam, J.; Adler, S.; Agarwal, S.; Ahmad, L.; Akkaya, I.; Aleman, F.L.; Almeida, D.; Altschmidt, J.; Altman, S.; Anadkat, S. Gpt-4 technical report. *arXiv* **2023**, arXiv:2303.08774. [\[CrossRef\]](#)
61. Team, G.; Anil, R.; Borgeaud, S.; Alayrac, J.-B.; Yu, J.; Soricut, R.; Schalkwyk, J.; Dai, A.M.; Hauth, A.; Millican, K. Gemini: A family of highly capable multimodal models. *arXiv* **2023**, arXiv:2312.11805. [\[CrossRef\]](#)
62. Team, K.; Du, A.; Yin, B.; Xing, B.; Qu, B.; Wang, B.; Chen, C.; Zhang, C.; Du, C.; Wei, C. Kimi-vl technical report. *arXiv* **2025**, arXiv:2504.07491. [\[CrossRef\]](#)
63. Cappelli, M.A.; Di Marzo Serugendo, G. Methodological Exploration of Ontology Generation with a Dedicated Large Language Model. *Electronics* **2025**, *14*, 2863. [\[CrossRef\]](#)
64. Jiang, L.; Shi, J.; Wang, C.; Pan, Z. Intelligent control of building fire protection system using digital twins and semantic web technologies. *Autom. Constr.* **2023**, *147*, 104728. [\[CrossRef\]](#)
65. Krishnasamy, L.; C, S.; Dhanaraj, R.K.; Al-Khasawneh, M.A.; Al-Shehari, T.; Alsadhan, N.A.; Selvarajan, S. Intelligent traffic congestion forecasting using BiLSTM and adaptive secretary bird optimizer for sustainable urban transportation. *Sci. Rep.* **2025**, *15*, 18423. [\[CrossRef\]](#) [\[PubMed\]](#)
66. Horrocks, I.; Patel-Schneider, P.F.; Boley, H.; Tabet, S.; Grosz, B.; Dean, M. SWRL: A semantic web rule language combining OWL and RuleML. *W3C Memb. Submiss.* **2004**, *21*, 1–31.
67. Ye, P.; Zhang, C.; Du, J.; Zhang, X. Domain Knowledge Aggregation Model Based on Knowledge Graph and Large Language Model. In Proceedings of the 2025 5th International Conference on Consumer Electronics and Computer Engineering (ICCECE), Dongguan, China, 28 February–2 March 2025; pp. 586–589.
68. Tupayachi, J.; Xu, H.; Omitaomu, O.A.; Camur, M.C.; Sharmin, A.; Li, X. Towards next-generation urban decision support systems through ai-powered construction of scientific ontology using large language models—A case in optimizing intermodal freight transportation. *Smart Cities* **2024**, *7*, 2392–2421. [\[CrossRef\]](#)

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.