Ontology-based multi-source heterogeneous O&M data integration framework for tunnel structural health assessment

Longxiang Liu[a], Peng An[a], Zhengru Ren[a], Ru An[b], Jiarui Lin[c], Yutao Guo[a], Zhen-Zhong Hu[a, *]

[a] Institute for Ocean Engineering, Shenzhen International Graduate School, Tsinghua University, Shenzhen, 518055, China.

[b] Shenzhen Urban Transport Planning Center Co., Ltd, Shenzhen, 518057, China

[c] Department of Civil Engineering, Tsinghua University, Beijing, 100084, China

[*] Corresponding author

Abstract: As a typical representative of infrastructure, tunnels are indispensable carriers for the normal operation of cities, with their safe and efficient operation directly influencing urban efficiency. However, the various data supporting tunnel operation and maintenance (O&M) exhibit significant diverse sources and structural differences, which pose substantial challenges to tasks such as tunnel structural health assessment. To address these challenges, this paper proposes an ontology-based multi-source heterogeneous O&M data integration framework to support the assessment of tunnel structural health, thereby improving decision-making efficiency in tunnel maintenance. The framework consists of four layers: data layer, ontology layer, mapping layer, and application layer, enabling the unified modeling, integration, and comprehensive application of multi-source heterogeneous tunnel O&M data. Additionally, the proposed framework is applied to a practical engineering project, the Tanglang Mountain Tunnel. Compared with existing methods, the framework demonstrates improvements in data fusion accuracy, data completeness, and operational efficiency.

Keywords: Ontology; Fuzzy comprehensive evaluation; Semantic web technology; Structural health assessment

1. Introduction

With the continuous advancement of urbanization, the scale of tunnels has been expanding, gradually becoming an indispensable component for the normal operation of cities. However, compared to the vast scale of tunnel construction, the level of tunnel O&M management remains relatively low, resulting in significant deficiencies in both safety and efficiency. Issues such as the diversity of O&M data sources and their low levels of integrated utilization expose critical gaps in tunnel O&M management.

The O&M process of tunnels generates amounts of data resources, including various models retained during the design phase, monitoring data for structural health, as well as resource data such as documents, images, and videos accumulated during the maintenance phase. These data are often stored in different databases, file systems, or hardware devices, originating from diverse sources and structural differences, which pose significant challenges for tunnel O&M management (Leng et al., 2020). Therefore, one key issue in the tunnel O&M process is how to integrate and utilize these heterogeneous data to support applications such as tunnel structural health assessment.

Currently, numerous methodologies exist for the integration and fusion of heterogeneous data. Traditional data fusion methods primarily focus on the integration of data from multiple sensors, including adaptive weighting approaches (Pan et al., 2020), Bayesian methods (Wang et al., 2018; Yoon & Yu, 2017), Kalman filtering (Ma et al., 2022), fuzzy theory (Zhang et al., 2014), and Dempster-Shafer evidence theory (Denœux, 2016). While these methods address the heterogeneity issues among similar data sources, their effectiveness diminishes when applied to data with significant structural and format differences. Specifically, traditional data fusion methods are unable to extract semantic information from multi-source heterogeneous data and lack the ability to infer new knowledge from existing data, thereby impeding the achievement of more profound semantic fusion.

With the rapid development of the semantic web (Pauwels et al., 2017) and knowledge engineering (Kügler et al., 2023), the concept of ontology has entered the forefront of attention. Ontology, owing to its semantic consistency, robust data integration capabilities, and data inference functionalities, presents a more precise, consistent, and intelligent approach to handling and querying data. Garnering widespread attention from researchers, ontology demonstrates significant advantages in integrating semantic information from heterogeneous data. Originally a concept in philosophy, ontology was later introduced into the realms of artificial intelligence and information science. In 1993, Gruber from Stanford University's Knowledge Systems Laboratory defined ontology as "an explicit specification or representation of a conceptualization" (Gruber, 1993). It describes knowledge at the semantic level, serving as a universal conceptual model for domain knowledge (Pardo et al., 2012). By establishing ontologies for various heterogeneous data (Le & David Jeong, 2016; Venugopal et al., 2015), semantic information interchange and fusion among data with different structures can be achieved (Mignard & Nicolle, 2014).

Against the above backdrop, this paper presents an ontology-based multi-source heterogeneous O&M data integration framework to support tunnel structural health assessment, thereby facilitating tunnel maintenance. The framework comprises four layers: (1) Data layer: Data collected during tunnel O&M processes is categorized, with methods for associating each category with the ontology provided. (2) Ontology layer: Various approaches are employed to develop ontologies for different types of O&M data. (3) Mapping layer: Ontology mapping is used to establish relationships between the different ontologies. (4) Application layer: Based on the tunnel O&M data fusion model, the structural health status of the tunnel structure is assessed. Additionally, a practical engineering project is utilized to validate the framework's effectiveness.

The main contributions of this paper are as follows:

- Various ontologies are developed for multi-source heterogeneous O&M data in tunnel maintenance. Compared to previous studies, these ontologies incorporate a broader spectrum of data types relevant to the tunnel O&M phase and are directly associated with real-world data, thereby achieving unification at both the data and ontology layers.
- An ontology mapping method is proposed based on concept similarity and local confidence,

which calculates similarity from different aspects. By comprehensively considering the different semantic features of the ontologies, the mapping relationships are established, ultimately forming a cohesive integration model for heterogeneous data.

- Building on the integrated tunnel O&M data model, a tunnel structural health assessment strategy is proposed. This strategy leverages ontologies to extract various types of tunnel O&M data and employs a comprehensive approach to assess the structural health of tunnels, thereby mitigating the potential decision-making errors associated with reliance on single-source data.

The remainder of this study is organized as follows. In section 2, the relevant research on ontology modeling, ontology mapping, and tunnel structural health assessment is introduced. Section 3 presents the proposed ontology framework, along with its detailed components. Section 4 validates the framework through a case study, providing a comparison with other methods and a discussion of the results. Finally, section 5 concludes the study while elaborating on future work.

2. Literature review

2.1 Ontology modeling in infrastructure domain

Nowadays, numerous ontologies tailored to various heterogeneous data have been developed, including IfcOWL ontology for Building Information Modeling (BIM) data (Pauwels & Terkaj, 2016), SSN for monitoring data, and Building Topology Ontology (BOT) for architectural topological information (Rasmussen et al., 2017). In addition to the development of ontologies for individual data types, many scholars are now focusing on ontologies for various data types in the infrastructure domain to facilitate compelling correlation and interaction among heterogeneous data. One approach is to develop a reference ontology. For instance, Deng et al. (2016) developed a reference ontology called Semantic City Model, which includes all entities and attributes of the BIM data standard Industry Foundation Classes (IFC) and the Geographic Information System (GIS) data standard City Geographic Markup Language (CityGML). By establishing mapping relationships between the Semantic City Model and the original schema, effective integration of BIM and GIS can be achieved. In the field of building fire protection, Liu et al. (2023) developed an ontology model for building fire protection (BFP), organizing, classifying, and connecting various entities from four aspects: system, device, operation, and construction. This model serves as a bridge linking the geometric information of buildings with sensor data, thereby integrating geometric information with fire monitoring sensor information and providing data support for subsequent fire protection systems.

Another approach involves associating and mapping concepts from different data domain ontologies to develop a comprehensive ontology for heterogeneous data fusion. For instance, Hor et al. (2016) developed corresponding ontologies based on the attributes and relationships of IFC and CityGML. They then utilized ontology mapping methods to identify similar concepts and relationships, thereby achieving the integration of BIM and GIS ontology models. Shi et al. (2023) developed an ontology for City Information Modeling (CIM) to integrate BIM, CIM, and Internet of Things (IoT) data. This ontology comprises two parts: initially developing an ontology for BIM-GIS integration, followed by associating dynamic IoT monitoring data with it, thereby forming a comprehensive city information model.

The previous methods have facilitated the effective integration of heterogeneous data ontologies. However, several challenges persist in the application of ontologies. Currently, there is a lack of relevant O&M data ontologies in the tunnel domain, and the methods for ontology modeling of different O&M data types in tunnels lack systematic research. Additionally, most current ontology models involve simplistic data modeling, lacking in-depth comprehension and application capability.

2.2 Ontology mapping methods and systems

In the field of data fusion, the integration of large amounts of multi-source heterogeneous data is often involved, and a single ontology is generally insufficient to cover various heterogeneous data types. Therefore, establishing effective associations between different O&M data ontologies through ontology mapping is considered a practical approach. Numerous studies have applied various ontology mapping methods to integrate ontologies from different domains. Based on research analysis, related studies can be categorized into four types:

(1) Ontology semantic similarity matching method: This approach compares the semantic similarity of different ontology concepts to identify the relationships between heterogeneous ontologies. OntoDNA is an ontology mapping system focused on dynamically addressing semantic inconsistencies between ontologies by treating ontology concepts as DNA sequences for similarity calculation (Kiu & Lee, 2006).

(2) Ontology structural similarity matching method: This method analyzes the structural relationships between different ontologies to discover existing mapping rules. For instance, Huang et al. (2020) proposed a semantic processing-based ontology structural similarity calculation method, which measures the similarity between different concept nodes based on their semantic distance in the ontology. Falcon combined lexical similarity and structural similarity matching strategies, identifying mapping relationships by analyzing class and attribute names, definitions, and their structural relationships in the ontology (W. Hu & Qu, 2008).

(3) Ontology instance matching method: This approach utilizes the instance data of ontologies and algorithms such as machine learning to find mapping relationships between ontologies. GLUE is an instance-based ontology mapping system that applies machine learning to discover mapping relationships between ontologies, evaluating the similarity between concepts from multiple perspectives to achieve more accurate mapping results (Doan, 2002).

(4) Comprehensive method: These methods integrate the above approaches to identify mapping relationships between ontologies. ASMOV is an ontology mapping system that combines multiple similarity measurement methods, including text similarity, structural similarity, and semantic similarity, to assess the similarity between concepts in two ontologies from multiple angles (Jean-Mary et al., 2009). RiMOM is an ontology mapping system based on Bayesian decision theory, which transforms the ontology mapping discovery problem into a minimal risk decision problem (J. Li et al., 2009). It

incorporates multiple strategies, including string-based, structural, and semantic-based methods, to improve mapping accuracy from different dimensions.

This paper draws on the strengths of previous ontology mapping methods and systems, considering ontology similarity at different levels, including concepts, instances, and structures. It dynamically adjusts the weights of different strategies based on the characteristics of different ontologies, thereby improving the accuracy and stability of the alignment between various O&M data ontologies.

2.3 Tunnel structural health assessment

Due to the diversity of monitoring data collected from operational tunnels, it is necessary to establish a comprehensive health indicator evaluation system to assess tunnel structural health status. In the late 20th century, Einstein et al. (1995) first introduced risk assessment theory into tunnel engineering. They integrated various construction uncertainties and environmental uncertainties into consideration, establishing tunnel cost models and tunnel risk decision support systems. This provided theoretical guidance for tunnel risk assessment research. Subsequently, Kampmann et al. (1998) further proposed a classification system for tunnel engineering based on this foundation. Using the Copenhagen Metro project as a case study, they employed the Monte Carlo method to establish a tunnel risk assessment model and conducted qualitative analysis on the likelihood of accidents. However, during this period, tunnel health assessment methods primarily relied on qualitative research, lacking the introduction of quantitative evaluation indicators. Such approaches no longer suffice to meet the needs of personnel for tunnel structural health assessment.

The analytic hierarchy process (AHP) offers a solution for quantifying the weights of different indicators. AHP, a quantitative analysis method for complex decision-making problems, was proposed by American scholar Thomas Saaty. It has been applied in some tunnel health assessments, mitigating the influence of subjective factors to some extent (Hyun et al., 2015). However, its reliance on domain experts to rate different indicators limits its practical application in engineering projects. To address this limitation, fuzzy theory has been introduced into tunnel structural health assessment. By calculating the membership degrees of different indicators for various health levels, fuzzy theory provides a more objective reflection of actual health conditions. For instance, Khademi Hamidi et al. (2010) utilized expert surveys and fuzzy AHP to propose solutions for risk management in tunnel design, construction, and operation using the Resalat Tunnel as a background. Zhang et al. (2014), based on a comprehensive evaluation model using fuzzy AHP, integrated different types of sensor data into the health grading of shield tunnels to assess tunnel safety conditions. Ren et al. (2023) designed a five-level evaluation index system according to common sensor layouts in shield tunnels, establishing corresponding multi-level health evaluation factor sets to assess the structural health of different tunnel monitoring locations. Sun et al. (2015) proposed a specialized fuzzy AHP comprehensive evaluation model, constructing a six-level indicator evaluation system for health monitoring data. These studies, incorporating fuzzy theory, have achieved graded assessment of tunnel structural health, becoming commonly used evaluation strategies today.

Building upon the previous research, this study simplifies the hierarchical division of tunnels by incorporating commonly observed monitor indicators, facilitating practical applications in engineering. Additionally, by developing ontologies for heterogeneous tunnel O&M data, this paper integrates heterogeneous data to assess the health status of the tunnel structure, thereby aiding in the selection of subsequent tunnel O&M strategies.

3. The proposed framework based on ontology

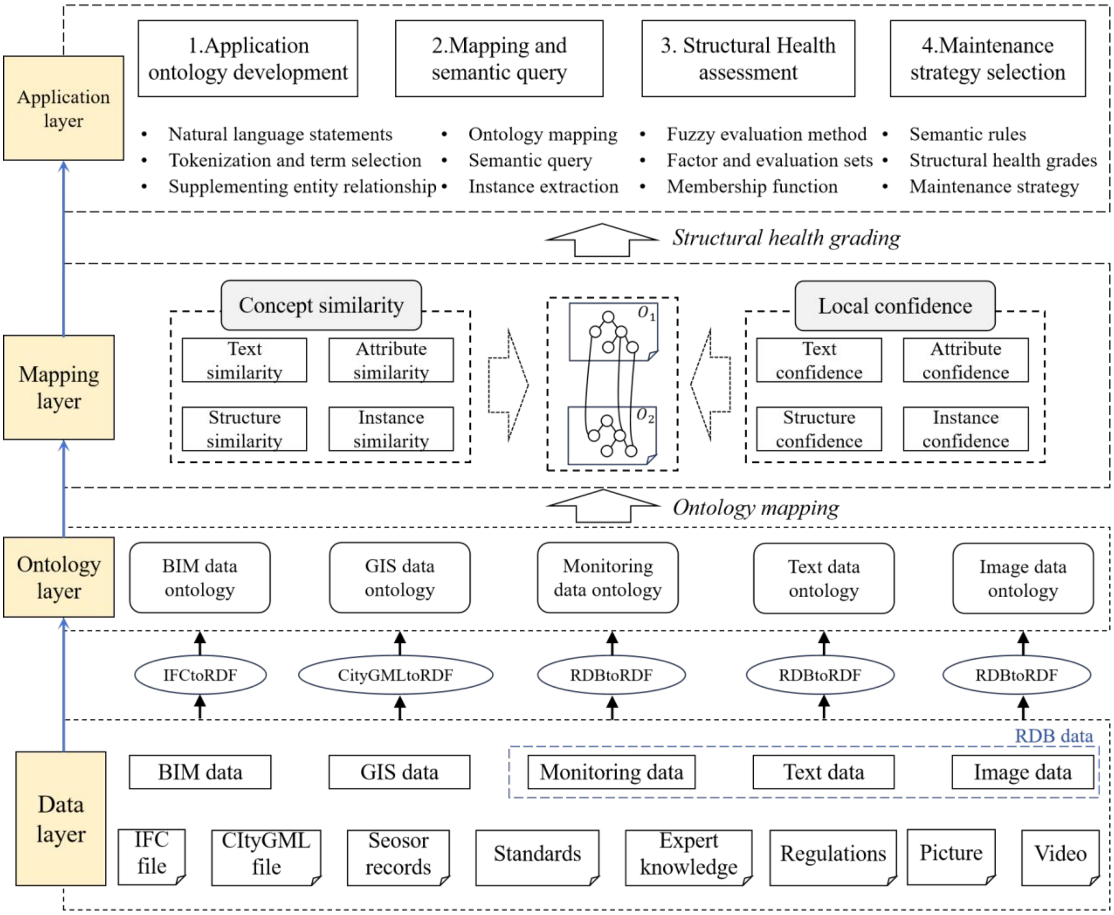


Figure 1. Overall ontology-based framework.

To integrate and utilize various data for tunnel structural health assessment, this paper proposes a framework based on ontologies for tunnel heterogeneous information modeling, ontology mapping, and structural health assessment. As illustrated in Figure 1, the overall framework consists of four layers: data layer, ontology layer, mapping layer, and application layer. The data layer encompasses the commonly used data in tunnel O&M processes, including BIM data, GIS data, monitoring data, text data, and image data. The ontology layer primarily focuses on the conceptual definitions and developing methods of the BIM data ontology, GIS data ontology, monitoring data ontology, text data ontology, and image data ontology. The ontology layer is linked to the data layer through various transformation methods. The mapping layer calculates the comprehensive similarity between different ontologies using concept similarity and local confidence, thereby establishing connections among different ontologies. The application layer leverages the obtained tunnel heterogeneous O&M data ontologies to classify the structural health status of the tunnel and inform subsequent maintenance strategies.

3.1 Data layer

The data generated during tunnel O&M processes originates from diverse sources and exhibits significant structural variability. To facilitate the effective integration and utilization of heterogeneous O&M data, it is essential to establish a classification for tunnel O&M data. This classification will provide a foundation and data support for subsequent unified ontology modeling. Currently, the primary sources of tunnel O&M data include BIM models, GIS models, monitoring data, image data, video surveillance, inspection reports, maintenance logs, and regulatory standards. These data encompass the majority of information required throughout the tunnel O&M process. Aside from BIM and GIS models, most of the other O&M data can be stored in relational databases (RDB), which allows them to be categorized as RDB files. Consequently, this paper classifies tunnel heterogeneous O&M data into three main categories: BIM data, GIS data, and RDB data. Based on data formats, RDB data can be further subdivided into monitoring data, text data, and image data. Moreover, given the varying degrees of tunnel complexity, the types of O&M data differ accordingly, allowing users to flexibly extend the data as needed. These actual data can be associated with the ontologies developed later in this paper through methods such as IFCtoRDF, direct mapping, D2RQ and so on.

3.2 Ontology layer

Ontologies are typically categorized into top-level, domain, task, and application ontologies based on their hierarchy and dependencies. This paper focuses on developing a domain ontology for tunnel facilities to describe relevant concepts, properties, and axioms. Various methods exist for ontology development, including the seven-step method (Noy & McGuinness, 2001), Methontology (Fernández-López et al., 1997), skeleton-based methods (Alfaifi, 2022), IDEF5 (L. Li et al., 2021), and so forth. Among above methods, the seven-step method is commonly used, with specific steps including 1) determining the domain and scope of the ontology, 2) considering reusing existing ontologies, 3) enumerating important terms in the ontology, 4) defining the classes and the class hierarchy, 5) defining the properties of classes, 6) defining the facets of the properties, and 7) creating instances. This paper adopts the seven-step method for ontology development, and efforts are made to reuse existing entities and relationships to maintain consistency with current ontologies.

3.2.1 BIM data ontology

The latest version of the IFC standard is IFC 4.3.2.0, which introduces concepts including *Ifc:Road*, *Ifc:Railway*, *Ifc:MarineFacility*, *Ifc:Bridge*, and *Ifc:Building* for the interpretation and description of specific infrastructure domains. In IFC, these concepts are grouped under *Ifc:Facility* within *Ifc:SpatialStructureElement*. A spatial structure element is a generalization of all spatial elements used to define a spatial structure, providing spatial organization for a project. *Ifc:FacilityPart* allows for the spatial breakdown of built facilities. Currently, IFC has not introduced entities and properties specifically related to tunnels. To develop the ontology for tunnel BIM data, this paper extends the concepts and relationships of tunnel facilities based on existing IFC entity concepts (Venugopal et al., 2015). The core content is depicted in Figure 2.

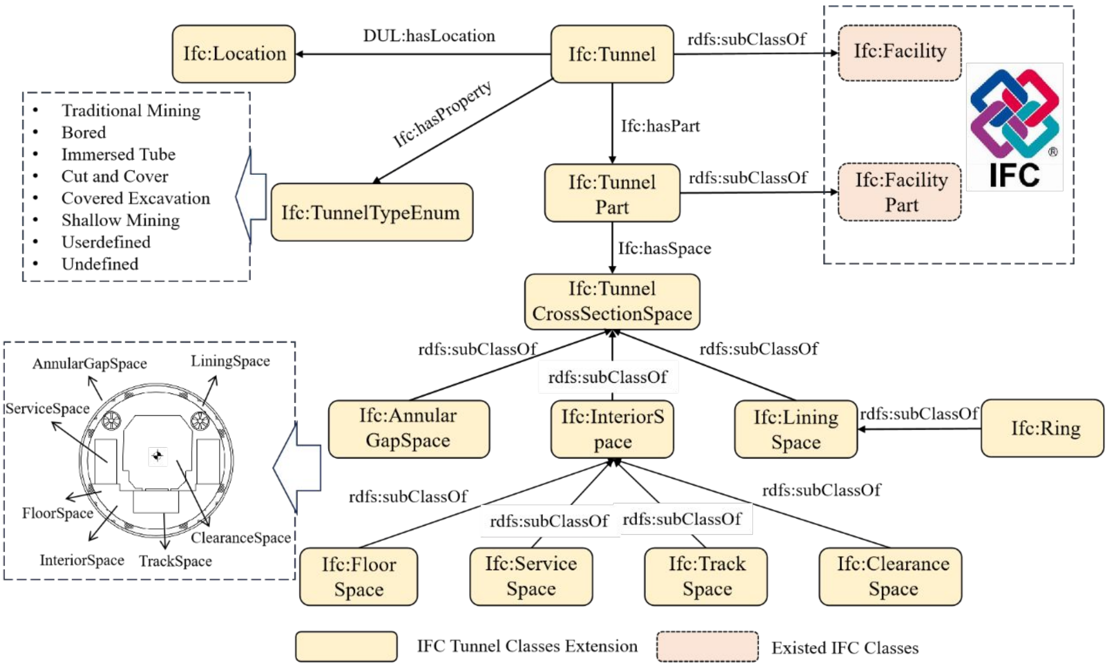


Figure 2. Tunnel BIM data ontology.

3.2.2 GIS data ontology

To develop a tunnel GIS data ontology based on CityGML, this paper extracts relevant modules from the CityGML standard, primarily including the Core module, the Construction module, and the Tunnel module. These core concepts are further expanded and refined. Taking the core concept ontology as an example, the relevant hierarchies and semantic relationships are illustrated in Figure 3. The core concept ontology primarily provides definitions for object concepts, spatial concepts, and geometric concepts, facilitating the modeling of data pertaining to the tunnel's surrounding environment (Z.-Z. Hu et al., 2019). The construction ontology serves as a higher-level concept for the tunnel ontology, offering definitions for basic elements and spatial concepts, while the tunnel ontology specifically defines the relevant entities associated with tunnel concepts.

3.2.3 Monitoring data ontology

SSN serves as an ontology for describing sensor networks. Its objective is to provide a unified semantic description for sensor networks, facilitating better understanding and integration of sensor data. SSN follows a modular architecture, both horizontally and vertically, and comprises a lightweight core ontology known as sensor, observation, sample, and actuator (SOSA). SOSA is an ontology for describing sensor networks, observation data, and actuators. Both SOSA and SSN are based on open standards and semantic technologies such as OWL and RDF. They have gained widespread application and support in knowledge engineering, often used together to model and describe sensor networks and related entities. Based on the concepts and relationships of SSN and SOSA, the ontology for tunnel monitoring data is developed. The detailed concepts and relationships are depicted in Figure 4.

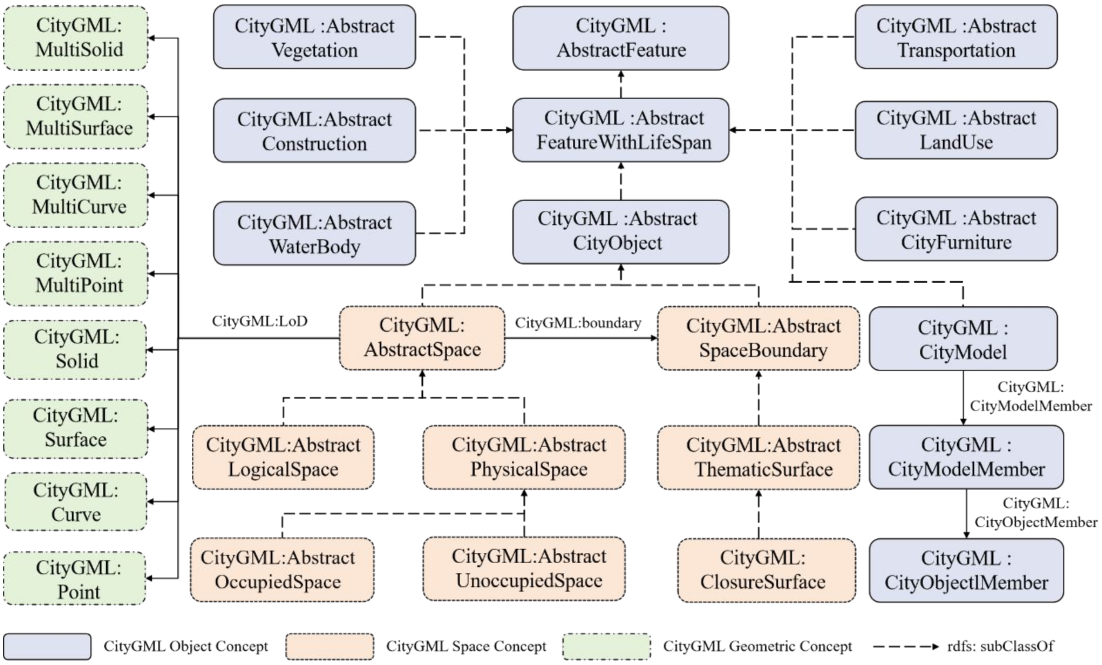


Figure 3. Tunnel GIS core concept ontology

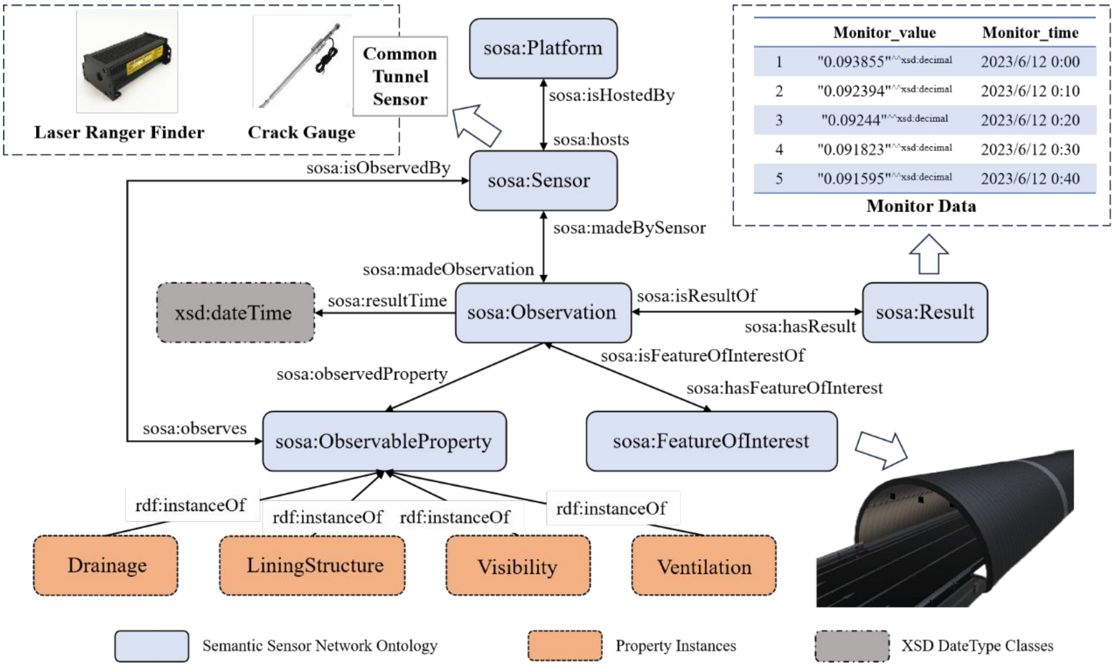


Figure 4. Tunnel monitoring data ontology.

3.2.4 *Text data ontology*

The text data generated during the actual O&M processes of tunnels typically includes inspection reports, maintenance logs, and other documents. These data are generally stored in databases; however, the design of these databases often varies significantly among different tunnel projects, making it challenging to establish a universal text data ontology applicable to all tunnels. Therefore, this paper focuses on methods for developing a text data ontology.

Relational databases (such as MySQL and SQL Server) store data in the form of tables, with each table containing a set of related data that typically represents distinct objects. As the fundamental data structure within a database, a table consists of rows (records) and columns (fields). Each row corresponds to a specific entity instance, while each column represents an attribute of that entity. Moreover, the table's attributes commonly include primary keys and foreign keys. A primary key is one or more fields within the table that uniquely identifies each row, making it indispensable; a foreign key, on the other hand, is a field used to establish a relationship between two tables. Based on the above analysis, this paper proposes the following ontology development rules to facilitate the development of the text data ontology: 1) The primary key of the data table is mapped to the core concept (class) of the ontology. 2) The primary key serves as a core concept to establish connections with other fields. 3) Other fields are mapped as object properties or data properties based on their data types. 4) Foreign keys connect different data tables. 5) Records within the data table are mapped as instances.

3.2.5 Image data ontology

The image data generated during tunnel O&M processes includes inspection images and surveillance videos. The storage of these data typically requires a combined application of relational databases and file systems. Relational databases (such as MySQL and PostgreSQL) can be utilized to store the metadata of images and videos, while the actual image and video files are stored within the file system. The methodology for developing the ontology of image data is analogous to that of text data.

3.3 Mapping layer

Section 3.1 and 3.2 have addressed the instantiation of various types of tunnel O&M data and the development of their corresponding ontologies. This section primarily focuses on establishing associations among the ontologies of different domains through ontology mapping. Ontology mapping is a complex process. Based on the research of Ehrig et al. (2004), ontology mapping can be divided into six steps: ontology feature extraction, candidate entity pair selection, similarity calculation, similarity integration, similarity interpretation, and iterative computation. Building on this foundation, this paper further refines the process, proposing eight steps that include ontology input, ontology preprocessing, candidate entity pair selection, concept similarity calculation, comprehensive similarity calculation, mapping result integration, iterative computation, and ontology output. Among these, concept similarity calculation and comprehensive similarity calculation are core steps in the ontology mapping process, as illustrated in Figure 5.

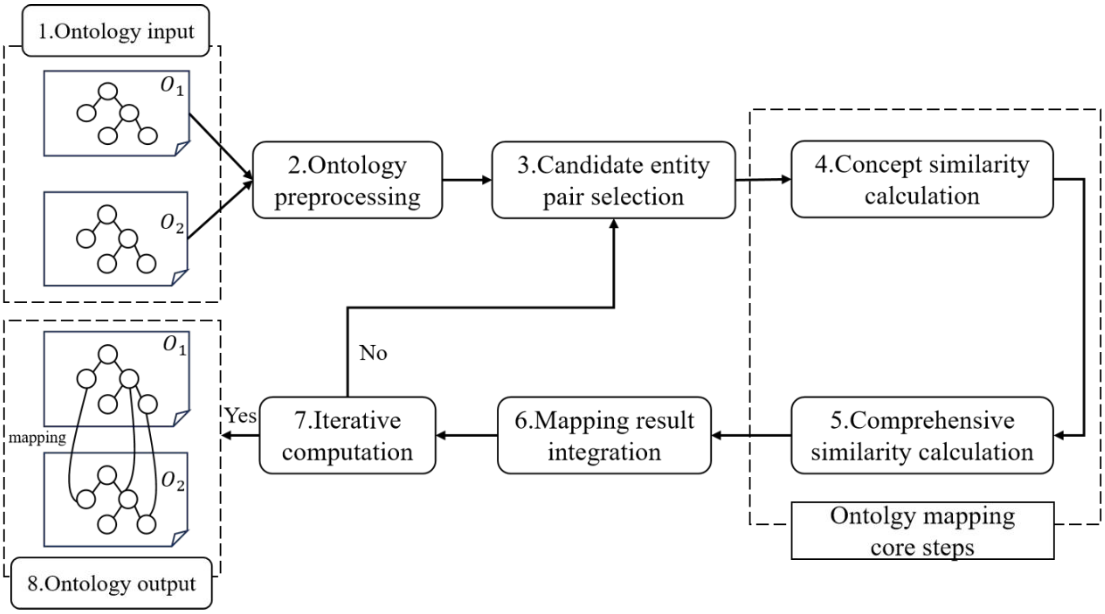


Figure 5. Ontology mapping flowchart

(1) Ontology input

Determine the ontologies O_1 and O_2 that need to be mapped based on the various tunnel O&M data. These ontologies typically adhere to standard formats and encompass the entity concepts, attributes, and instances required for subsequent mapping.

(2) Ontology preprocessing

Prior to mapping, ontologies often require a series of preprocessing operations to ensure that their structure and format are consistent and standardized. These preprocessing steps typically include the normalization of entity naming conventions, unification of format conversions, elimination of duplicate entity concepts, and simplification of complex conceptual hierarchies.

(3) Candidate entity pair selection

Upon completing the preprocessing operations, the next step is to traverse and retrieve the concepts from the ontologies. This paper primarily focuses on entity concept pairs $\{(c_1,c_2)|c_1 \in O_1, c_2 \in O_2\}$ between two ontologies, wherein only one candidate entity pair is selected during a single ontology mapping process.

(4) Concept similarity calculation

The next step involves calculating the similarity for the selected candidate entity pair. There are various methods for computing concept similarity; in this paper, it is focused on calculating text similarity, attribute similarity, structural similarity, and instance similarity.

(a) Text-based concept similarity

Edit distance is a string metric that measures the difference between two string sequences. It represents the minimum number of operations required to transform one string into another through insertions, deletions, and substitutions. The smaller the edit distance, the higher the similarity between the two strings. The detailed calculation formula is shown in Equation (1):

$$Sim_{string}(c_1,c_2) = 1 - \frac{dis(c_1,c_2)}{max\ (|c_1|,|c_2|)} \tag{1}$$

Here, $|c_1|$ and $|c_2|$ represent the string lengths of two entity concepts. $dis(c_1,c_2)$ represents the edit distance between the two entity concepts c_1 and c_2, which is the total number of operations required to transform one string into the other.

(b) Attribute-based concept similarity

The similarity is calculated based on the shared attributes of the concept nodes, with the specific calculation formulas given in Equations (2) and (3):

$$Sim_{property}(c_1,c_2) = \frac{f(c_1 \cap c_2)}{f(c_1 \cup c_2) - \lambda f(c_1 - c_2) - (1-\lambda) f(c_2 - c_1)} \tag{2}$$

$$\lambda = \begin{cases} \dfrac{f(c_1 - c_2)}{f(c_1 \cup c_2) - f(c_1 \cap c_2)} & f(c_1) \geq f(c_2) \\ 1 - \dfrac{f(c_1 - c_2)}{f(c_1 \cup c_2) - f(c_1 \cap c_2)} & f(c_1) < f(c_2) \\ 1 & f(c_1 \cup c_2) = f(c_1 \cap c_2) \end{cases} \tag{3}$$

Here, $f(c_1)$ and $f(c_2)$ represents the number of attributes of concept c_1 and c_2; $f(c_1 \cap c_2)$ represents the number of common attributes between concepts c_1 and c_2; $f(c_1 \cup c_2)$ denotes the total number of attributes of concepts c_1 and c_2; $f(c_1 - c_2)$ represents the number of attributes that concept c_1 possesses but concept c_2 does not; and $f(c_2 - c_1)$ represents the number of attributes that concept c_2 possesses but concept c_1 does not.

(c) Structural-based concept similarity

The similarity is determined by calculating the geometric distance between two concepts, specifically the distance from each concept node to the nearest common subnode, which can then be used to determine the semantic distance. It is generally considered that the smaller the semantic distance, the higher the similarity between the two concepts. The specific calculation is shown in Equation (4):

$$Sim_{structure}(c_1,c_2) = \frac{Dep(c|O_1) + Dep(c|O_2)}{(dis(c_1,c_2) + 1) \times (Dep(O_1) + Dep(O_2))} \tag{4}$$

Where the definitions of $dis(c_1,c_2)$ are as follows.

$$dis(c_1,c_2) = L(c_1,c) + L(c_2,c) \tag{5}$$

Here, c is the nearest common parent node of c_1 and c_2, $L(c_1,c)$ represents the shortest path from concept c_1 to c, and $L(c_2,c)$ represents the shortest path from concept c_2 to c. In addition to considering the distance between the concept nodes and the common parent node, this formula also takes into account the depth of the ontology and the depth of the common parent node in the ontology, denoted as $Dep(c|O_1)$, which represents the depth of the common parent node in the ontology O_1, and $Dep(O_1)$, which represents the depth of the ontology O_1.

(d) Instance-based concept similarity

Given two sets A and B, the Jaccard coefficient is defined as the ratio of the size of the intersection of A and B to the size of their union, as defined in Equation (6):

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \tag{6}$$

When sets A and B are empty, $J(A,B) = 1$. Based on the Jaccard calculation formula, the instance-based concept similarity calculation is given by Equation (7):

$$Sim_{instance}(c_1,c_2) = \frac{f(|I_{c_1} \cap I_{c_2}|)}{f(|I_{c_1} \cup I_{c_2}|)} \tag{7}$$

Here, I_{c_1} represents the instance set corresponding to concept c_1, I_{c_2} represents the instance set corresponding to concept c_2, $f(|I_{c_1} \cap I_{c_2}|)$ denotes the number of elements in the intersection of I_{c_1} and I_{c_2}, and $f(|I_{c_1} \cup I_{c_2}|)$ represents the number of elements in the union of I_{c_1} and I_{c_2}.

(5) Comprehensive similarity calculation

In step (4), the similarity of multiple candidate entity pairs is computed. To integrate these similarities, this step is based on the concept of local confidence, where local confidences are calculated for text, attribute, structural, and instance similarities. In existing methods for integrating similarity, a fixed weight approach is typically used, which ignores the applicability of different strategies to different ontologies. By introducing local confidence, dynamic combinations of multiple mapping strategies can be achieved, resulting in more reliable mapping results.

(a) Text-based local confidence

In the calculation of text-based local confidence, the Term Frequency-Inverse Document Frequency (TF-IDF) method is used. First, the classes and annotated properties of the two ontologies to be mapped are concatenated to construct corresponding documents. Then, the TF-IDF method is applied to vectorize the document content corresponding to the ontologies. Finally, the cosine similarity is used to calculate the similarity between the two vectors. The specific calculation formula is given in Equation (8):

$$\alpha_{string} = \frac{TFIDF(D_1) \cdot TFIDF(D_2)}{|TFIDF(D_1)| \times |TFIDF(D_2)|} \tag{8}$$

Here, D_1 and D_2 are the two documents corresponding to the ontology O_1 and O_2. $TFIDF(D_1)$ and $TFIDF(D_2)$ are the feature vectors obtained after transforming the documents using the TF-IDF method. $|TFIDF(D_1)|$ and $|TFIDF(D_2)|$ represent the norms of the feature vectors. Through this calculation, the overall textual similarity between the two ontologies can be measured.

(b) Attribute-based local confidence

The calculation formula for attribute-based local confidence is given in Equation (9):

$$\alpha_{property} = \frac{Common(|Prop|)_{O_1,O_2}}{min\ (Prop_{O_1},Prop_{O_2})} \tag{9}$$

Here, $Common(|Prop|)_{O_1,O_2}$ represents the number of shared attributes between the two ontologies, and $\min\ (Prop_{O_1},Prop_{O_2})$ represents the smaller of the number of attributes in the two ontologies. This formula determines the local confidence of attribute similarity by measuring the proportion of shared attributes in relation to the total number of attributes in the two ontologies.

(c) Structural-based local confidence

In the calculation of structural-based local confidence, the following formula is used:

$$\alpha_{structure} = \frac{Same(|Dep|)_{O_1,O_2} \cap Same(|\mathrm{Sub}|)_{O_1,O_2}}{\min\ (C_{o_1},C_{o_2})} \tag{10}$$

Here, $Same(|Dep|)_{O_1,O_2} \cap Same(|\mathrm{Sub}|)_{O_1,O_2}$ represents the number of concepts in both ontologies that have the same depth and number of sub-concepts, and $\min(C_{o_1}, C_{o_2})$ represents the smaller of the number of concepts in the two ontologies. This formula primarily reflects the structural differences between the two ontologies.

(d) Instance-based local confidence

In the calculation of instance-based local confidence, the following formula is provided:

$$\alpha_{instance} = \left(1 - \frac{a}{|I_{O_1}| + |I_{O_2}| + a}\right) \times \frac{\min\{|I_{O_1}|, |I_{O_2}|\}}{\max\{|I_{O_1}|, |I_{O_2}|\}} \tag{11}$$

Here, $|I_{O_1}|$ and $|I_{O_2}|$ represent the number of instances for ontology O_1 and O_2, respectively, and a is a tuning factor to ensure proper scaling, typically set to a constant value, $a = 1$. From the formula, it can be seen that instance-based local confidence is positively correlated with the sum of the number of instances in the mapped ontologies, and negatively correlated with the difference in the number of instances between the mapped ontologies.

In practice, when calculating the local confidence under different strategies, a confidence threshold is often preset. When the calculated local confidence is below this threshold, the confidence of the similarity calculated by that strategy is considered low and can be approximated as 0. Ultimately, this process yields a comprehensive similarity score for the candidate entity concept pairs. The calculation formula is presented as follows.

$$\mathrm{Sim}(c_1, c_2) = \frac{\sum_k \alpha_k \times Sim_k}{\sum_k \alpha_k} \tag{12}$$

Where $k \in \{string, property, structure, instance\}$, α_k, Sim_k represent the local confidence and concept similarity based on the k strategy, respectively.

(6) Mapping result integration

Select an appropriate similarity threshold to integrate candidate entity pairs that meet or exceed this threshold into the ontology, thereby establishing equality relationships between the two concepts. During the integration process, it is essential to ensure the correctness and consistency of the mapping results. In cases of conflict, manual corrections will be necessary.

(7) Iterative computation

After the calculation for the candidate entity pair is completed, the process returns to step (3) to select the next candidate entity pair. This cycle continues until all candidate entity pairs have been traversed, at which point the computation concludes.

(8) Ontology output

Output the final integrated target ontology. This output should encompass all mapping results while maintaining consistency and completeness. The resulting ontology can be utilized for further data integration, querying, and analysis.

3.4 Application layer

3.4.1 Overview of tunnel structural health grading strategy

Building on the previously discussed data layer, ontology layer, and mapping layer, this section proposes a structural health grading strategy for tunnels. The strategy comprises several steps: 1) Application ontology development. This step involves O&M personnel inputting corresponding requirement statements into the tunnel multi-source heterogeneous O&M data management platform to develop the relevant application ontology based on O&M needs. 2) Mapping and semantic query. Utilizing the ontology mapping methods in the mapping layer, the developed application ontology can be associated with the tunnel O&M data ontology. This mapping allows for the extraction of concepts from different O&M data ontologies through semantic query, thereby facilitating access to the required tunnel O&M data from diverse sources and structures. 3) Tunnel structural health assessment. Using fuzzy comprehensive evaluation methods, the associated multi-source heterogeneous O&M data is subjected to comprehensive analysis and computation, thereby obtaining structural health assessment values for the tunnel at different levels. 4) Maintenance method selection. Referencing relevant regulatory standards, an assessment of the tunnel structural health grading is conducted. Based on the tunnel structural health grades, appropriate O&M strategies are selected.

With the core objective of this strategy centered on the assessment of structural health status, the effective utilization of integrated multi-source heterogeneous data necessitates a systematic and structured evaluation approach. To this end, the present study establishes a four-level hierarchical evaluation index system, comprising the following levels: 1) specific tunnel monitoring indicators (including stress, displacement, crack width, electromechanical status, etc.), 2) tunnel cross-sections, 3) tunnel segments, and 4) the overall tunnel structure. Concurrently, in accordance with the Technical Specifications for Highway Tunnel Maintenance (JTG H12-2015), the structural safety condition is categorized into five distinct grades: Grade 1 (Intact condition), Grade 2 (Slight damage), Grade 3 (Moderate damage), Grade 4 (Severe damage), and Grade 5 (Dangerous condition). To enable quantitative assessment, a fuzzy comprehensive evaluation method is employed, wherein expert-defined threshold intervals are integrated with weightings derived through AHP method. This methodological framework facilitates the fusion of tunnel O&M data into a unified, quantifiable health index for each tunnel section.

3.4.2 Application ontology development

O&M personnel can input various natural language statements into the tunnel multi-source heterogeneous O&M data management platform according to operational needs, thereby developing different application ontologies. The specific construction process is illustrated in Figure 6. Initially, a tokenization tool is employed to segment the input natural language statements, resulting in the corresponding tokenized sequence of statements. Next, several core terms are identified, and the entity relationships among core terms are established. Finally, the corresponding application ontology is generated. O&M personnel may develop corresponding application-specific ontologies based on actual requirements, thereby enabling the retrieval of targeted tunnel O&M data.

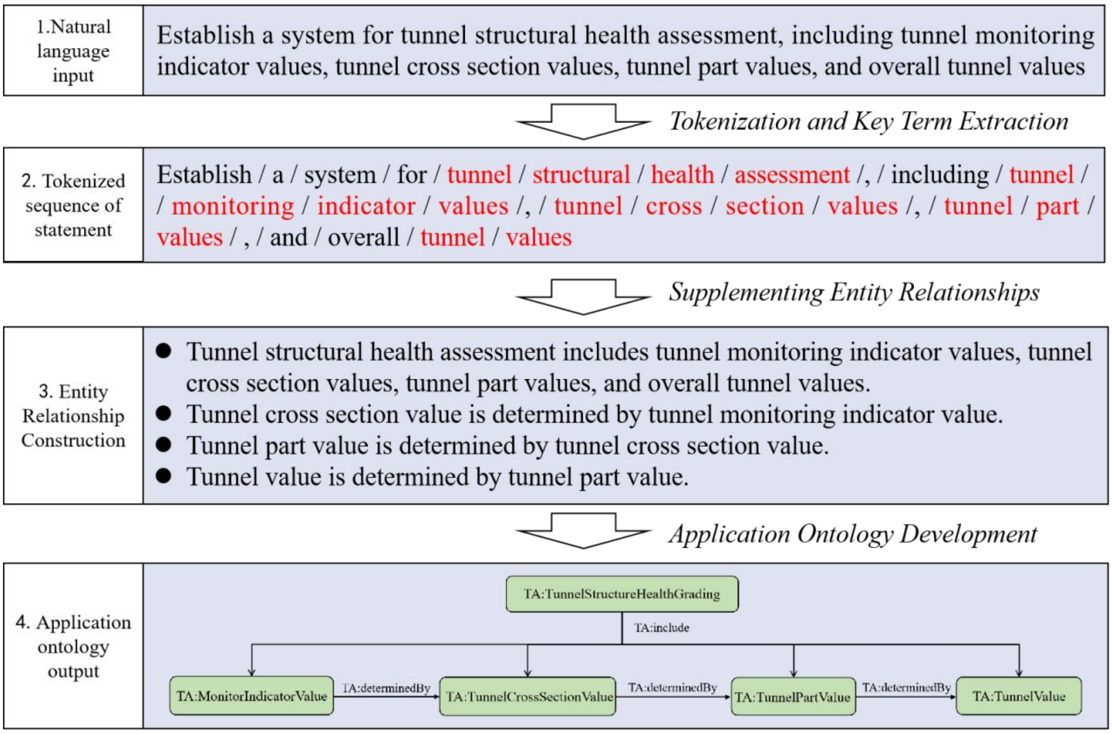


Figure 6. Application ontology development flowchart

Using the tunnel structural health assessment as an example, it is essential to assess and classify the structural health of different segments and cross-sections of the tunnel. Based on the review of relevant literature on tunnel structural health assessment in Section 2, it is evident that the evaluation of tunnel structural health requires the establishment of a multi-tiered system of health indicators. This system generally includes levels such as tunnel monitoring indicators, specific tunnel components, tunnel cross-sections, and tunnel segments. In this paper, a four-tier structural health assessment ontology is developed, as illustrated in Figure 6. This ontology is prefixed with "TA" and encompasses structural health evaluation information across various levels of the tunnel. Specifically, *TA:TunnelValue, TA:TunnelPartValue, TA:TunnelCrossSectionValue, TA:MonitorIndicatorValue* represent the overall structural health assessment value of the tunnel, the structural health assessment value of specific monitored segments, the structural health assessment value of tunnel cross-sections, and the structural health assessment value of a specific monitoring indicator, respectively. Among these assessment values, the higher-tier structural health assessment values are determined by the lower-tier ones, with the structural health assessment value of tunnel monitoring indicators being influenced by specific monitoring results, which include sensor data, text data, image data, and other types of heterogeneous O&M data for the tunnel.

3.4.3 *Mapping and semantic query*

After completing the development of the application ontology for tunnel structural health assessment, the application ontology is mapped to the tunnel multi-source heterogeneous O&M data ontology. This mapping enables the extraction of the required concepts and instances from the O&M data ontology. The ontology mapping method employed is based on the comprehensive similarity mapping method proposed in the mapping layer. Different types of O&M data provide varying information regarding tunnel structural health assessment. Monitoring data, text data and image data primarily offer specific structural health monitoring indicators for the tunnel. BIM and GIS data mainly provide detailed structural and locational information about the tunnel. Part of the ontology mapping is illustrated in Figure 7.

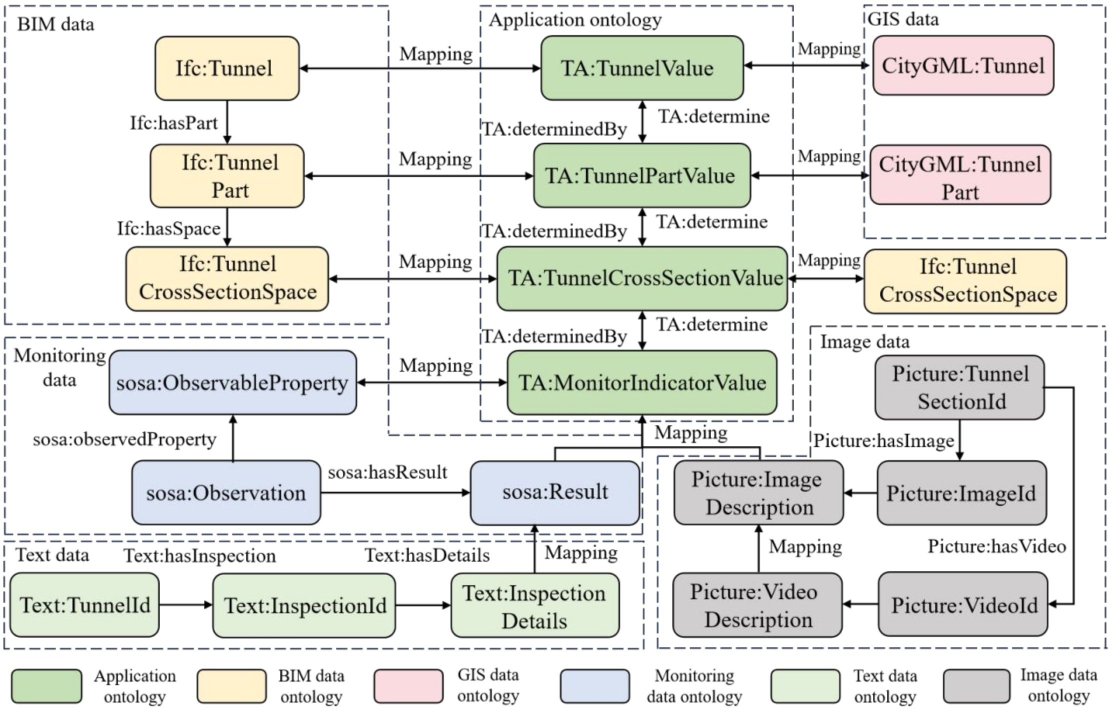


Figure 7. Mapping of application ontology and heterogeneous O&M data ontology

After obtaining the mapping results between the application ontology and the tunnel heterogeneous O&M data ontology, the required O&M data instances can be extracted using SPARQL Protocol and RDF Query Language (SPARQL) queries. SPARQL is a query language designed for querying and manipulating RDF data (Zhong et al., 2018). By importing ontology and instance data into GraphDB, SPARQL statements can be customized to achieve queries across heterogeneous information.

3.4.4 Structural health assessment

Upon obtaining the necessary heterogeneous O&M data for the tunnel, the next step involves a comprehensive utilization and analysis of these data to accurately assess the structural health status of the tunnel. This analysis lays the foundation for subsequent strategy selection in tunnel maintenance. In the acquired heterogeneous O&M data, BIM data primarily provides detailed structural and locational information about the tunnel. GIS data offers information about the surrounding environment and supplementary descriptions of relevant locations. Monitoring data encompasses long-term indicators such as stress, strain, and settlement; image data focuses on monitor indicators suitable for periodic inspections, such as tunnel cracks and structural damage; while text data primarily presents monitor indicators related to electromechanical equipment, traffic signs, and tunnel entrances in a text format.

The health values of monitor indicators reflected by the text and image data are determined by relevant O&M personnel. For the health values represented by the monitoring data, a fuzzy comprehensive evaluation method is predominantly employed for calculation. Additionally, BIM and GIS data enhance the information regarding the tunnel's cross-sections, segments, and locations by correlating with the O&M data. The overall methodological flowchart is illustrated in Figure 8.

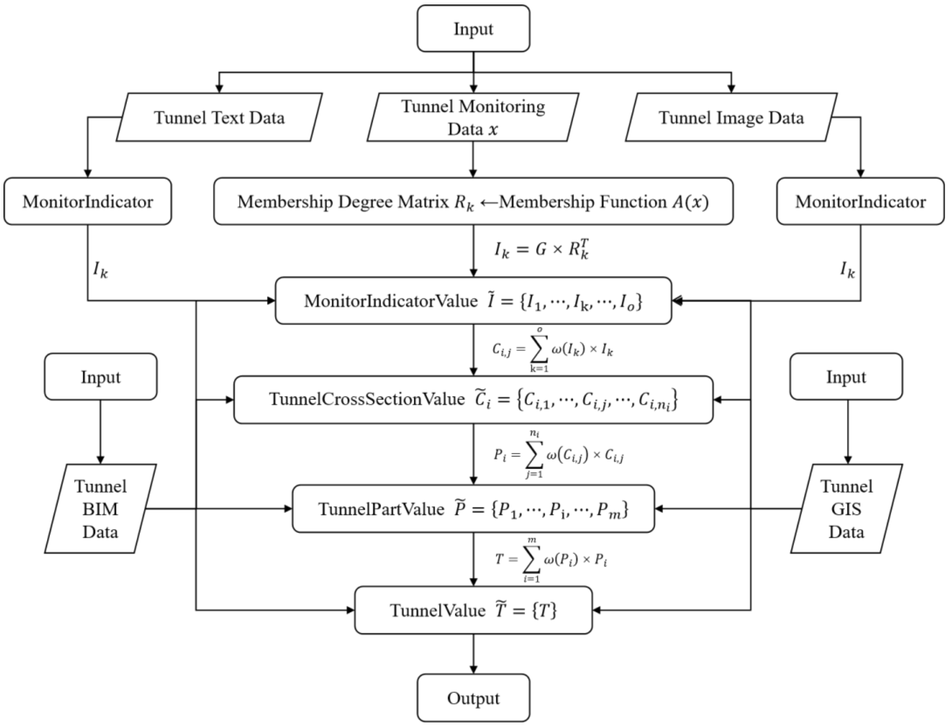


Figure 8. Overall flowchart of the tunnel structural health assessment strategy.

To assess the tunnel structural health based on fuzzy evaluation methods, corresponding factor sets for each hierarchical health assessment value are established. The relevant symbols, their corresponding relationships with tunnel structural health assessment ontology entities at various levels, and the specific meanings of each symbol are detailed in Table 1.

Table 1 The multi-level factor sets for tunnel structural health assessment.

Level	Ontology class	Factor set	Description
Ⅰ	TunnelValue	$\tilde{T} = \{T\}$	T represents the overall assessment value of the tunnel structural health.
Ⅱ	TunnelPartValue	$\tilde{P} = \{P_1,\cdots,P_i,\cdots,P_m\}$	P_i represents the structural health assessment value for the i-th section of the tunnel.
Ⅲ	TunnelCrossSectionValu e	$\tilde{C}_i = \{C_{i,1},\cdots,C_{i,j},\cdots,C_{i,n_i}\}$	$C_{i,j}$ represents the structural health assessment value for the j-th cross-section within the i-th monitoring section of the tunnel.
Ⅳ	MonitorIndicatorValue	$\tilde{I} = \{I_1,\cdots,I_k,\cdots,I_o\}$	I_k represents the health assessment value for the k-th monitor indicator within the tunnel.

In accordance with the tunnel structure evaluation method of Technical Specifications for Highway Tunnel Maintenance, the assessment of tunnel structural health is categorized into five grades: Grade 1, Grade 2, Grade 3, Grade 4, and Grade 5. These grades correspond to intact condition, slight damage, moderate damage, severe damage, and dangerous condition, respectively. Based on this classification, the evaluation set is defined as

$$\tilde{V} = \{v_1, v_2, v_3, v_4, v_5\} \tag{13}$$

where v_1, v_2, v_3, v_4, v_5 correspond to the previous structural health condition grades, respectively. To facilitate quantitative analysis of different grades, a grade scoring matrix is constructed as

$$G = [5,4,3,2,1] \tag{14}$$

where Grade 1 corresponds to 5, and the rest follow in descending order.

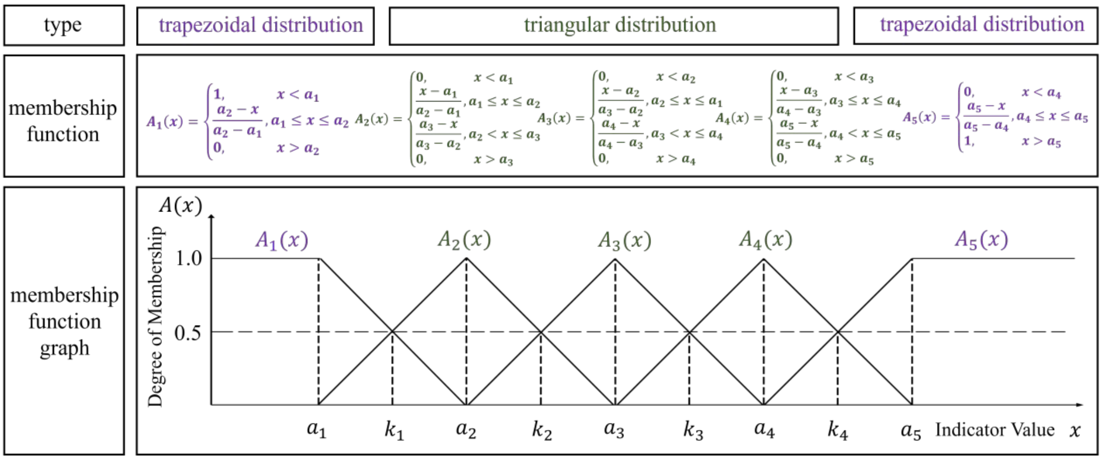


Figure 9. Membership distribution, function and graph.

In fuzzy comprehensive evaluation, membership functions are utilized to describe the membership relationship of actual data to the evaluation set (Lin et al., 2022). Common membership functions include rectangular, triangular, trapezoidal, parabolic, and normal distribution (T. H. Y. Li et al., 2013). In actual engineering, the health status of various monitoring factors generally conforms to a linear relationship with their monitoring values. Therefore, this paper employs a combination of triangular and trapezoidal distribution to construct a 5-grade membership function. The function is employed to describe the membership degree of actual data to different grades of structural health conditions. As illustrated in Figure 9, x represents the actual measurement of a monitoring indicator, and k_1, k_2, k_3, k_4 are the classification boundaries for the monitoring indicator, which need to be determined based on the specific facility conditions. The relationships between k_1, k_2, k_3, k_4 and a_1, a_2, a_3, a_4, a_5 are as follows. Typically, a_1 is chosen to be $0.8k_1$, and the formulas for each membership function can be derived accordingly.

$$\begin{bmatrix} k_1 \\ k_2 \\ k_3 \\ k_4 \end{bmatrix} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0.5 & 0.5 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0 & 0.5 & 0.5 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \tag{15}$$

Based on the five grades defined in the evaluation set, a fuzzy comprehensive judgment matrix is constructed. For a monitoring indicator I_k corresponding to the set of monitor indicators $\tilde{I}$, the membership degree matrix for the actual measurement x of the monitoring indicator is given by

$$R_k = [A_1(\mathrm{x}), A_2(\mathrm{x}), A_3(\mathrm{x}), A_4(\mathrm{x}), A_5(\mathrm{x})] \tag{16}$$

The health assessment value of this monitoring indicator is calculated by

$$I_k = G \times R_k^T \tag{17}$$

The health values I_k of monitor indicators reflected by the text and image data are determined by relevant O&M personnel and are collectively categorized within $\tilde{I}$. The calculation of the tunnel health assessment values for each cross-section, section, and the overall structure, is illustrated in Table 2, where $\omega(P_i)$ represents the weight value of each section of the tunnel, generally determined based on the length of the tunnel section; $\omega(C_i)$ represents the weight of each cross-section of the tunnel, unless specified, it is generally assumed that the weight values of each cross-section are equal; $\omega(I_i)$ represents the weight of each monitor indicator, determined by AHP.

Table 2 The formula for calculating the multi-level structure health assessment value.

Level	Factor set	Formula
I	$\tilde{T} = \{T\}$	$T = \sum_{i=1}^{m} \omega(P_i) \times P_i$
II	$\tilde{P} = \{P_1,\cdots,P_i,\cdots,P_m\}$	$P_i = \sum_{j=1}^{n_i} \omega(C_{i,j}) \times C_{i,j}$
III	$\tilde{C}_i = \{C_{i,1},\cdots,C_{i,j},\cdots,C_{i,n_i}\}$	$C_{i,j} = \sum_{k=1}^{o} \omega(I_k) \times I_k$

3.4.5 Maintenance strategy selection

After calculating the health assessment values for various levels of tunnel structure, the obtained values are associated with the developed tunnel structural health assessment ontology. Subsequently, a series of specific semantic rules for tunnel structural health grades are defined to achieve a graded assessment of tunnel structural health. Using tunnel overall structural health assessment values as an example, the following rule can be derived based on the fuzzy comprehensive evaluation matrix constructed in Section 3.4.3. Here, *TA:hasRate* is an object property representing the structural health grade of the tunnel entity.

$$\left.\begin{array}{r} lessThan(TunnelValue,1.5) \\ greaterThanOrEqual(TunnelValue,1) \end{array}\right\} => hasRate(Tunnel,5)$$

By defining semantic rules for the structural health of different tunnel levels, structural health grades can be obtained for various monitoring indicators, monitoring cross-sections, tunnel segments, and the overall tunnel. Once the structural health grade of a tunnel has been determined, appropriate and targeted maintenance strategies can be deployed. In alignment with relevant standards, a systematic mapping is established between structural health grades and the corresponding O&M interventions, as detailed in Table 3. Each grade reflects a distinct physical condition of the tunnel structure and necessitates a tailored response. 1) Grade 1: The overall tunnel structure is in sound condition, with no abnormalities or only minor, non-progressive anomalies. No immediate action is required; routine inspections should continue at the standard frequency. 2) Grade 2: The tunnel structure exhibits minor deterioration but remains stable. Regular monitoring of the affected areas is advised, including inspections of crack width, joint integrity, and electromechanical system performance. 3) Grade 3: Moderate structural damage is present, though deterioration progresses slowly. Localized reinforcement is recommended, including crack injection, waterproof lining repairs, and grouting in zones exhibiting elevated strain. 4) Grade 4: Severe structural deterioration is observed, necessitating immediate maintenance action. Reinforcement measures may involve the installation of steel ribs, fiber-reinforced polymer (FRP) wrapping, or shotcrete application. 5)

Grade 5: The tunnel is in a critically compromised state, requiring emergency intervention. Full closure is warranted, followed by urgent structural assessment and reinforcement, which may include partial reconstruction to restore structural integrity.

Table 3 Maintenance strategies for varied tunnel structural health grades.

Grade	Value range	Evaluation set	Description	Maintenance strategy
Grade 1	[4.5,5]	v_1	Intact condition	Routine maintenance
Grade 2	[3.5,4.5)	v_2	Slight damage	Conduct surveillance on damaged structural segments and undertake necessary maintenance when essential.
Grade 3	[2.5,3.5)	v_3	Moderate damage	Conduct focused surveillance on compromised structural segments and implement localized maintenance and repairs as required.
Grade 4	[1.5,2.5)	v_4	Severe damage	Swiftly implement remedial measures for structural ailments
Grade 5	[1,1.5)	v_5	Dangerous condition	Promptly close the tunnel for necessary treatment. In exceptional circumstances, undertaking localized reconstruction or renovation may be imperative.

Following the selection of overall tunnel-level maintenance strategies, more granular measures must be implemented based on specific monitoring indicators corresponding to different tunnel segments and cross-sections. These localized indicators guide targeted interventions to ensure structural integrity and serviceability. For instance, when excessive cracking is detected in the tunnel lining, remedial actions such as epoxy injection, crack sealing, and renewal of waterproof linings should be undertaken. In cases of significant displacement or settlement, grouting reinforcement of the foundation and surrounding rock mass is necessary to restore structural stability. Should water leakage occur, corrective measures may include drainage system repairs and reinforcement of the waterproofing layer to mitigate moisture ingress. In the event of corrosion of metallic components, replacement with corrosion-resistant materials is recommended to prolong service life. For more severe structural deformations, comprehensive strengthening techniques such as steel rib support, FRP wrapping, and structural realignment should be employed to restore the tunnel's designed geometry and load-bearing capacity.

By linking the evaluation results with well-established repair and reinforcement methods, the framework supports precise, data-driven maintenance decisions. Additionally, health grades can be used to prioritize limited resources and plan long-term maintenance schedules.

4. Validation of the proposed framework

4.1 Case study

4.1.1 Case overview

The Tanglang Mountain Tunnel, completed in 2006, is situated in the Nanshan District of Shenzhen, traversing the Tanglang Mountain. This dual-tube six-lane highway tunnel is arranged separately for traffic traveling in opposite directions. Using Nanshan District towards Longgang District as the forward direction, determine the left and right tunnels accordingly. The axial distance between the left and right tunnels is 38 meters, with the left tunnel spanning a total length of 1719.5 meters and the right tunnel spanning 1711 meters. Taking into account various factors such as adverse geological conditions, structural defects, and safety hazards at different locations along the tunnel, the tunnel has been subdivided into distinct segments. Each segment prioritizes different monitoring parameters, with several monitoring sections selected within each segment for focused surveillance.

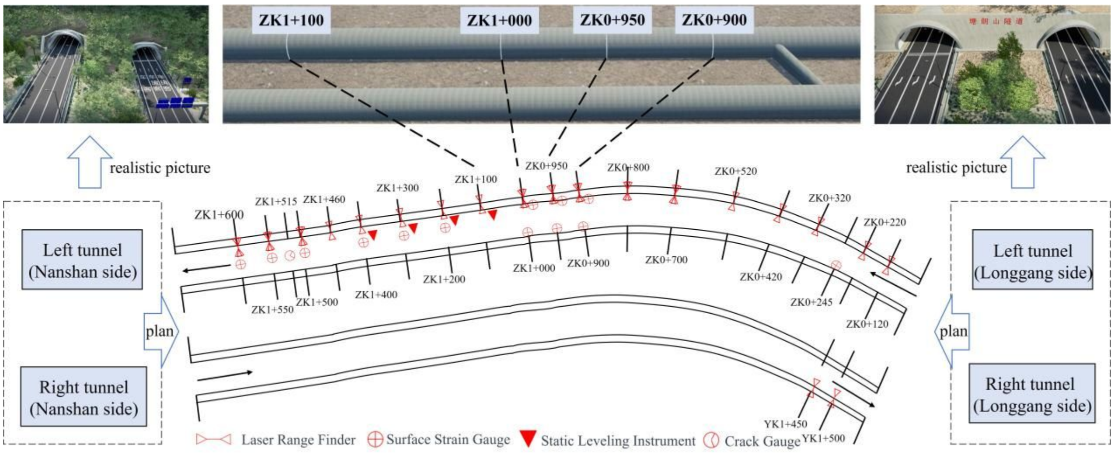


Figure 10. Pictures and sensor placements of Tanglang Mountain Tunnel.

Within the tunnel, four types of sensors are deployed: laser rangefinders, static leveling instruments, crack gauges, and surface strain gauges. These sensors facilitate real-time monitoring of various parameters. Different monitoring projects within the same monitoring section are arranged along the same cross-section, facilitating comparative analysis of monitoring data. Laser rangefinders are positioned at the right haunch of the monitoring section, while surface strain gauges are symmetrically placed at the crown of the arch and at both the left and right haunches of the monitoring section. Depending on site conditions, static leveling instruments are positioned at the haunches of the monitoring section. Crack gauges are conducted at the widest point of the crack, with one monitoring point installed. The detailed monitoring indicators, measurement ranges, and number of deployments for each sensor are shown in Table 4. This study primarily focuses on several segments of the left tunnel of the Tanglang Mountain Tunnel, with sensor placements illustrated in Figure 10.

Table 4 Basic information of sensors.

Sensor	Monitoring item	Range	Installation number
Laser Range Finder	Peripheral displacement	0.05~40m	28
Static Leveling Instrument	Tunnel settlement	0.2~100mm	4

Crack Gauge	Crack width	0~100mm	2
Surface Strain Gauge	Arch ring circumferential strain、arch crown longitudinal strain	0~3000$\mu\varepsilon$	12

4.1.2 Data instantiation and semantic query

According to Technical Guidelines for Long-term Monitoring of In-service Highway Tunnels (T/CHTS 10021-2020) and Technical Specifications for Highway Tunnel Maintenance (JTG H12-2015), considering the geological conditions, structural anomalies, and safety hazards at various locations within the tunnel, the tunnel has been segmented into distinct sections. Each section has specific focus areas for monitoring, and several monitoring cross-sections have been selected within each section for detailed scrutiny. This paper analyzes three key monitoring sections, with detailed information outlined in Table 5.

Table 5 Tunnel monitoring cross-section.

Monitoring part	Potential structural safety risk	Monitor indicator	Monitoring cross-section
ZK0+900~ZK1+000	Tunnel Eccentric Loading	Peripheral displacement, arch ring circumferential strain, electromechanical systems, crack numbers	ZK0+900 ZK0+950 ZK1+000
ZK1+100~ZK1+400	Differential Settlement and Structural Shrinkage	Peripheral displacement, tunnel settlement, arch crown longitudinal strain, electromechanical systems, crack numbers	ZK1+100 ZK1+200 ZK1+300 ZK1+400
ZK1+500~ZK1+600	Relaxation Loading of Surrounding Rock	Peripheral displacement, crack width, arch crown longitudinal strain, electromechanical systems, crack numbers	ZK1+500 ZK1+515 ZK1+550 ZK1+600

The data from the Tanglang Mountain Tunnel during the O&M process is associated with the respective O&M data ontologies. The detailed maintenance data selected is illustrated in Table 6. Among these, the monitoring data from March 17, 2023 to June 13, 2023, is selected, reflecting structural monitoring indicators such as displacement, settlement, stress, and strain. The text data primarily indicates the daily operational status of the tunnel's electromechanical systems, while the image data mainly documents the number of cracks in the tunnel. BIM data provide information on the tunnel segments and cross-sectional locations. Utilizing the mapping results from the mapping layer, a multi-source heterogeneous data ontology encompassing all actual O&M data for the tunnel is ultimately obtained. Subsequently, through SPARQL queries and the mapping relationships between the application ontology and the tunnel heterogeneous O&M data ontology, the required heterogeneous O&M data can be queried.

Table 6 Overview of Tanglang Mountain tunnel maintenance data.

Data type	Source	Format	Size	Data and ontology association tool
BIM data	Design models	.ifc	226MB	IFCtoRDF
Monitoring data	Sensors	.csv	205MB	D2RQ

Text data	inspection reports, maintenance logs, regulatory standards.	.pdf	131MB	D2RQ
Image data	Inspection images、surveillance videos	.png	3.52GB	D2RQ

4.1.3 Structural health assessment

Following the tunnel structural health assessment value calculation method outlined in Section 3.4.4, the monitoring intervals for each health monitoring indicator are determined based on actual tunnel conditions and relevant literature (Dong et al., 2008). The membership functions for each indicator are determined and the corresponding fuzzy comprehensive judgment matrices are established. Multiplying these matrices by the grade scoring matrix $G = [5,4,3,2,1]$ yields the structural health assessment values for monitoring indicators.

Table 7 Health monitoring intervals for partial monitoring indicators.

Grade	Peripheral displacement (mm)	Tunnel settlement (mm)	Crack width (mm)	Arch ring circumferential strain ($\mu\varepsilon$)	Arch crown longitudinal strain ($\mu\varepsilon$)
Grade 1	[0,5]	[0,5]	[0,0.1]	[0,50]	[0,50]
Grade 2	[5,15]	[5,10]	[0.1,0.5]	[50,100]	[50,100]
Grade 3	[15,30]	[10,20]	[0.5,1.0]	[100,200]	[100,200]
Grade 4	[30,60]	[20,50]	[1.0,5.0]	[200,400]	[200,400]
Grade 5	[60, +∞]	[50, +∞]	[5.0, +∞]	[400, +∞]	[400, +∞]

After calculating the structural health assessment values for different indicators, it is also necessary to establish the weights for each indicator. This study, referencing relevant literature (Zhong et al., 2018), utilizes the AHP to determine the weights of each indicator. Through consultation with relevant tunnel experts and engineers, the monitoring indicators involved in this engineering project are compared in terms of their weights. The exponential scale method $e^{\frac{0}{4}}$ to $e^{\frac{8}{4}}$ is used to construct the judgment matrix. Taking ZK1+200 as an example, three monitoring indicators are set as follows: peripheral displacement (I_1), tunnel settlement (I_2), and vault longitudinal strain (I_3). The judgment matrix is constructed as shown in the equation (18).

$$J = \begin{bmatrix} 1 & e^{-\frac{2}{4}} & e^{\frac{2}{4}} \\ e^{\frac{2}{4}} & 1 & e^{\frac{4}{4}} \\ e^{-\frac{2}{4}} & e^{-\frac{4}{4}} & 1 \end{bmatrix} \tag{18}$$

After normalization calculations, the weight vector is obtained as shown in equation (19).

$$W = \begin{bmatrix} 0.307 \\ 0.506 \\ 0.186 \end{bmatrix} \tag{19}$$

The consistency index (CI) is calculated by

$$CI = \frac{\lambda_{max} - n}{n-1} \approx 0.006 \tag{20}$$

The consistency rate (CR) is calculated by

$$CR = \frac{CI}{RI} \approx 0.010 < 0.1 \tag{21}$$

Since CR < 0.1, the consistency ratio is less than 0.1, indicating that the consistency of the judgment matrix is acceptable. Therefore, the calculation results of the structural health assessment value for the ZK1+200 cross-section are shown in Table 8.

Table 8 Structural health assessment value for cross-section ZK1+200.

$C_{i,j}$	$\omega(I_k)$	I_k
4.45	0.307	5.00
	0.506	3.93
	0.186	5.00

Table 8 indicates that the assessment values for three indicators of the cross-section ZK1+200 are as follows. Peripheral displacement, tunnel settlement, and arch crown longitudinal strain are rated at 5.00, 3.93, and 5.00. After constructing the judgment matrix using AHP, the final weight values for the three indicators are determined. The cross-section structural health assessment value for ZK1+200 is calculated to be 4.45. The computation process for other cross-sections is similar to the previous procedure.

By performing a weighted summation of the health assessment values of each monitoring indicator and their respective weights, the structural health assessment values can be ultimately derived for the corresponding cross-sections. And then the mean value of these cross-sections within the same section is computed to represent the structural health assessment value for that specific section. For distinct sections, the calculation of the tunnel's overall structural health assessment value is conducted based on the lengths of the sections.

Table 9 Structural health assessment values for various hierarchical levels of the Tanglang Mountain Tunnel.

T	$\omega(P_i)$	P_i	$\omega(C_{i,j})$	$C_{i,j}$	Monitor cross-section
4.43	0.17	4.20	0.33	3.69	ZK0+900
			0.33	5.00	ZK0+950
			0.33	3.92	ZK1+000
	0.66	4.47	0.25	4.41	ZK1+100
			0.25	4.45	ZK1+200
			0.25	4.52	ZK1+300
			0.25	4.51	ZK1+400
	0.17	4.48	0.25	5.00	ZK1+500
			0.25	4.35	ZK1+515
			0.25	4.93	ZK1+550
			0.25	3.65	ZK1+600

Table 9 shows that the structural health assessment values for the three selected tunnel sections are 4.20, 4.47, and 4.48, respectively. The overall structural health assessment value of the tunnel is 4.43, indicating that the overall structural health condition of the tunnel is good. However, there are also some cross-sections with relatively low structural health assessment values, such as ZK0+900, ZK1+000, and ZK1+600. These sections require specific remediation measures based on the monitoring indicators.

4.1.4 *Maintenance strategy selection*

Semantic web rule language (SWRL) is a language designed for representing rules on the semantic web (Z.-Z. Hu et al., 2022). It empowers users to define rules based on OWL ontology, facilitating inference and queries in knowledge representation. SWRL rules permit the assertion of logical relationships within the ontology, thus supporting more advanced semantic reasoning (Chen & Luo, 2019).

Utilizing SWRLTab plugin in protégé, inference can be achieved for the health grades of different hierarchical levels within the tunnel. The entire reasoning process consists of four procedures, as shown in Figure 11. 1) The established reasoning rules are input into SWRLTab. 2) The ontology and the established rules are then transmitted to the reasoning engine, including 5 rules, 27 classes, 87 entities, and 815 axioms. 3) The reasoning engine is executed, resulting in 260 axioms inferred. 4) The inferred axioms are transmitted back to the ontology model, thereby achieving knowledge reasoning.

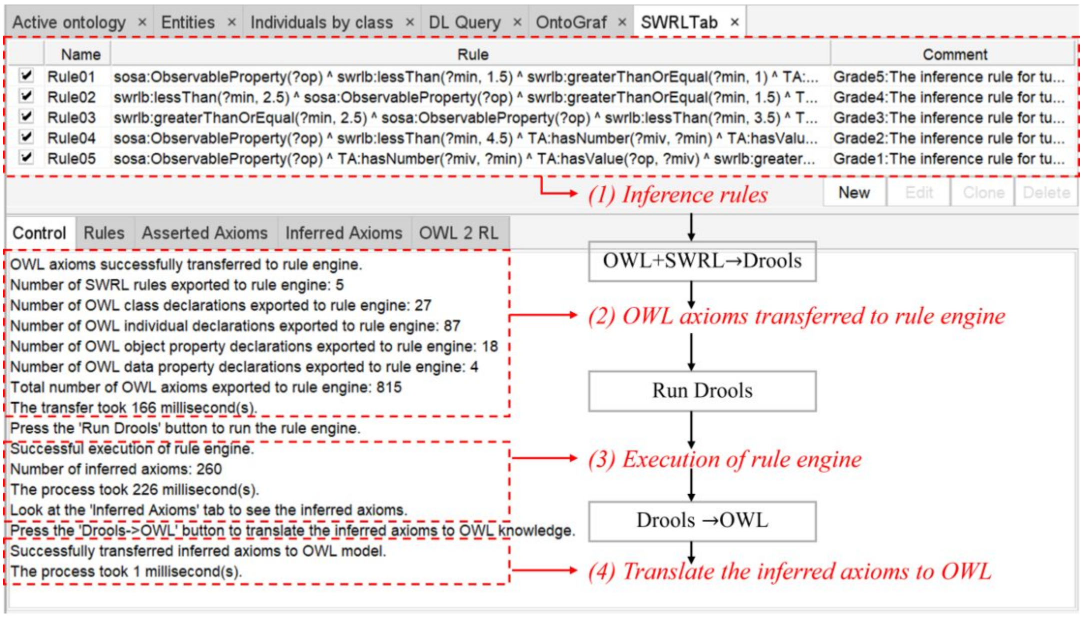
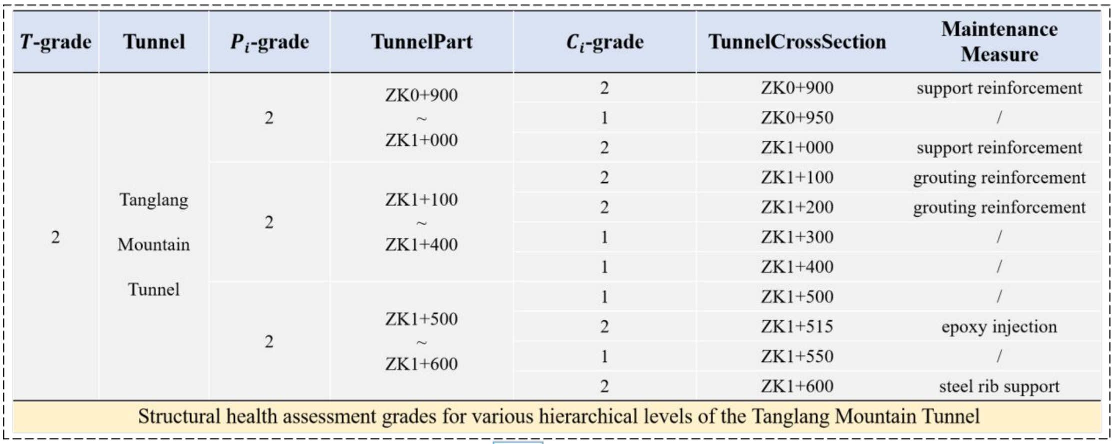


Figure 11. Reasoning process.

Through the above reasoning process, the structural health assessment grades of each level of the Tanglang Mountain Tunnel can be determined. As shown in Figure 12, the structural health assessment grades for the three selected tunnel sections are 2, 2, and 2, respectively. The overall structural health assessment grade of the Tanglang Mountain Tunnel is 2. Based on the current condition of the Tanglang Mountain Tunnel and according to the maintenance strategy of application layer, regular monitoring of the affected areas should be conducted, including inspections of crack width, joint integrity, and electromechanical system performance. With respect to specific tunnel monitoring indicators, consider the example of tunnel cross-section ZK1+200. The predominant structural concern at this location is excessive settlement, which may compromise structural performance if left unaddressed. Accordingly, targeted remediation measures such as grouting behind the lining and arch crown support installation should be considered in subsequent maintenance planning.



T-grade	Tunnel	P_i-grade	TunnelPart	C_i-grade	TunnelCrossSection	Maintenance Measure
2	Tanglang Mountain Tunnel	2	ZK0+900 ~ ZK1+000	2	ZK0+900	support reinforcement
				1	ZK0+950	/
				2	ZK1+000	support reinforcement
		2	ZK1+100 ~ ZK1+400	2	ZK1+100	grouting reinforcement
				2	ZK1+200	grouting reinforcement
				1	ZK1+300	/
				1	ZK1+400	/
		2	ZK1+500 ~ ZK1+600	1	ZK1+500	/
				2	ZK1+515	epoxy injection
				1	ZK1+550	/
				2	ZK1+600	steel rib support

Structural health assessment grades for various hierarchical levels of the Tanglang Mountain Tunnel

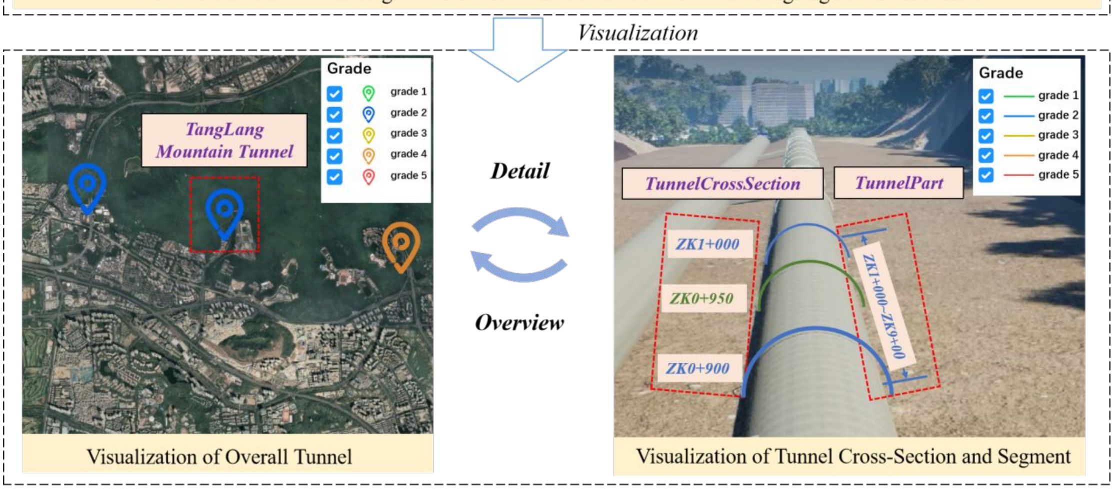


Figure 12. Structural health grade for Tanglang Mountain Tunnel.

4.2 Comparison with other methods

The proposed method is analyzed in terms of data fusion accuracy, completeness, and efficiency, and compare it with other methods:

Accuracy: The accuracy of data fusion in the proposed framework relies on whether the ontology mapping process can establish the relationships between different O&M data ontologies. To evaluate the accuracy of the ontology mapping, the standard benchmark from the Ontology Alignment Evaluation Initiative was selected for testing. The evaluation is carried out using parameters such as precision, recall, and F1 score, and the results are compared with traditional methods such as ASMOV, RiMOM, OntoDNA, and Falcon. As illustrated in Table 10, the proposed ontology mapping method performs better in terms of selected parameters, showing a certain level of improvement over previous methods.

Table 10 Comparison of ontology mapping methods

Algorithm	Precision	Recall	F1 score
ASMOV	0.92	0.87	0.89
RiMOM	0.92	0.88	0.90
OntoDNA	0.86	0.85	0.86
Falcon	0.92	0.88	0.90
The proposed framework	0.94	0.88	0.91

Completeness: Compared with previous knowledge-driven (Nuñez & Borsato, 2018) and data-driven approaches (Zhao et al., 2019), the framework encompasses a broader range of O&M data types, as illustrated in Table 11. By employing ontology for the unified modeling of various O&M data, it facilitates effective interactions among different data types, thereby reducing the potential decision-making errors that may arise from dependence on singular data sources.

Table 11 Comparison of involved data of different methods

Involved data	The proposed framework	The knowledge-driven method	The data-driven method
BIM data	√	√	
GIS data	√		
Monitoring data	√		√
Text data	√	√	
Image data	√		

Efficiency: Thanks to the semantic modeling and integration capabilities of ontologies, the proposed framework enables a fully automated process that encompasses data instantiation, semantic querying, structural health assessment, and decision-making. It achieves millisecond-level computation in data instantiation, mapping discovery, data extraction, and semantic reasoning, allowing for real-time monitoring of the tunnel structural health at all levels and significantly improving the efficiency of O&M decision-making. Especially in terms of data extraction, compared to traditional data-driven and manual search methods, the proposed framework shows significant improvements in speed. Some required time for framework function is shown in Table 12.

Table 12 Time required for each function within the framework

Framework function	Time	Unit
Data instantiation	83.5	ms per instance
Mapping discovery	92.3	ms per entity
Data extraction	16.67	ms per instance
Semantic reasoning	0.87	ms per axiom

4.3 Discussion

In the case section, this study specifically analyzes three segments of the Tanglang Mountain Tunnel. BIM data, monitoring data, text data, and image data collected during the tunnel's O&M processes are associated with the developed ontologies. The developed application ontology model is mapped to the tunnel O&M data ontologies, enabling the retrieval of various instances of multi-source heterogeneous O&M data through corresponding SPARQL queries. Utilizing the proposed tunnel structural health assessment strategy, these diverse O&M data are integrated and analyzed to obtain structural health assessment values for each level of the tunnel structure. Based on semantic reasoning rules and health assessment values, the structural health grading of the tunnel is automatically determined, assisting O&M personnel in selecting appropriate maintenance strategies. According to the results, the overall structural health assessment values of the selected three segments are all 2 or above, indicating a relatively good structural condition of the overall tunnel. This is consistent with the latest manual inspection report results, demonstrating the rationality of this strategy. For tunnel monitoring cross-sections with a structural health assessment grade of 2, targeted measures can be carried out for monitoring indicators with lower structural health assessment values.

The proposed framework exhibits strong scalability and generalizability, allowing for flexible extension across various dimensions, including types of O&M data, application ontology, monitoring indicators, and structural health assessment systems. This adaptability enables its application across tunnels of different scales and configurations. As an engineering case, the Tanglang Mountain Tunnel is employed to demonstrate the framework's applicability. The tunnel encompasses a wide array of real-world O&M data types and embodies most of the typical characteristics of conventional tunnels, thus serving as a representative case study. To further validate the robustness and versatility of the proposed framework, additional O&M datasets from diverse tunnel projects will be collected and analyzed in future work. These efforts aim to assess the framework's performance across multiple tunnel scenarios and support its continuous refinement and broader practical adoption.

Compared with existing methods, the proposed framework shows improvements in terms of accuracy, completeness, and efficiency of O&M data fusion. First, it comprehensively utilizes different similarity concepts such as text, attribute, structure, and instance, and introduces local confidence calculation formulas for each type of similarity, better establishing the relationships between different O&M data ontologies. Second, it integrates various heterogeneous data sources to assess the structural health of the tunnel, reducing potential decision-making errors that may arise from reliance on a single data source, thereby enhancing accuracy and comprehensiveness. Finally, it achieves millisecond-level fully automated computation and reasoning through steps such as data ontology instantiation, ontology mapping, data extraction, and semantic reasoning, significantly improving the efficiency of data fusion.

Although initial progress has been made in tunnel heterogeneous data integration and structural health assessment, it is undeniable that there are still several limitations in this study, which are summarized as follows.

（1）In establishing mapping relationships among different ontologies, this study primarily employs methods based on concept similarity and local confidence, yielding favorable outcomes. However, it is noteworthy that while this approach is effective for smaller ontologies, it becomes time-intensive and laborious when applied to large-scale projects with numerous ontology entities and attributes. Consequently, future plans include investigating automated ontology mapping to more efficiently establish mappings and associations across various ontologies.

（2）In tunnel structural health assessment, this study mainly employs a tunnel structural health assessment strategy based on fuzzy comprehensive evaluation. Although this method partially addresses the difficulty of quantitatively evaluating tunnel structural health, there is still subjectivity in the specific indicator level interval division and the determination of target layer object weights. Thus, further optimization is required in the future.

(3) The present study primarily focuses on periodic structural health assessment of tunnels and does not yet incorporate mechanisms for real-time response and maintenance. Future extensions of the proposed framework may explore the integration of real-time monitoring data through the deployment of edge computing nodes, enabling continuous acquisition and on-site processing of structural behavior metrics. By leveraging the expert-defined threshold boundaries established in this study, the system could incorporate an automated alert module, triggering maintenance recommendations when specific indicators exceed predefined limits. Additionally, AI-assisted decision-making may be employed to enhance predictive capabilities, enabling health trend forecasting and automated maintenance strategy recommendation.

5. Conclusion

The data generated during tunnel maintenance is both multi-source and structurally heterogeneous. By integrating heterogeneous data through unified modeling, the tunnel's structural health can be accurately assessed. This study proposes an ontology-based framework for tunnel O&M data integration, designed to support structural health assessment, thereby enhancing tunnel maintenance. This framework comprises four main layers: data layer, ontology layer, mapping layer, and application layer. Together, these layers form an integrated model for multi-source heterogeneous O&M data. The data layer focuses on categorizing tunnel O&M data and establishing linkages with the ontology layer. The ontology layer employs various methods to develop ontologies for each type of maintenance data. The mapping layer connects these ontology models through methods based on concept similarity and local confidence. The application layer encompasses application ontology development, mapping and semantic query, structural health assessment, and maintenance strategy selection. This approach substantially improves decision-making in tunnel maintenance.

In the validation section, the framework's efficacy is tested using the Tanglang Mountain Tunnel as a case study, followed by a comparative analysis with previous studies to highlight its advantages. To further enhance the framework’s accuracy, rationality, and applicability, the following future developments are planned: 1) Implementing an ontology automatic mapping method utilizing deep learning. Plans include leveraging advanced natural language processing (NLP) models to support the automatic mapping of entity concepts across different ontologies. 2) Extending the framework to additional infrastructures. This involves refining the tunnel structural health assessment strategies and broadening their application to include other critical infrastructures, such as bridges, highways, high-rise buildings, and beyond. 3) Future work will focus on expanding real-time monitoring and responsive maintenance capabilities, thereby promoting seamless integration of real-time O&M data.

Acknowledgements

This project is funded by the National Key Research and Development Program of China under Grant 2022YFC3801100.

References

Alfaifi, Y. (2022). Ontology Development Methodology: A Systematic Review and Case Study. *2022 2nd International Conference on Computing and Information Technology (ICCIT)*, 446–450. https://doi.org/10.1109/ICCIT52419.2022.9711664

Chen, H., & Luo, X. (2019). An automatic literature knowledge graph and reasoning network modeling framework based on ontology and natural language processing. *Advanced Engineering Informatics*, *42*, 100959. https://doi.org/10.1016/j.aei.2019.100959

Deng, Y., Cheng, J. C. P., & Anumba, C. (2016). Mapping between BIM and 3D GIS in different levels of detail using schema mediation and instance comparison. *Automation in Construction*, *67*, 1–21. https://doi.org/10.1016/j.autcon.2016.03.006

Denœux, T. (2016). 40 years of Dempster–Shafer theory. *International Journal of Approximate Reasoning*, *79*, 1–6. https://doi.org/10.1016/j.ijar.2016.07.010

Doan, A. (2002). *Learning to map between structured representations of data*. University of Washington.

Dong, Y., Xu, Y., Li, H., & Dai, M. (2008). A comparative study of the numerical scales and the prioritization methods in AHP. *European Journal of Operational Research*, *186*(1), 229–242. https://doi.org/10.1016/j.ejor.2007.01.044

Ehrig, M., & Staab, S. (2004). QOM – Quick Ontology Mapping. In S. A. McIlraith, D. Plexousakis, & F. Van Harmelen (Eds.), *The Semantic Web – ISWC 2004* (Vol. 3298, pp. 683–697). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-30475-3_47

Einstein, H., Chiaverio, F., & Koppel, U. (1995). Risk analysis for the Alder tunnel. *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, *4*(32), 190A.

Fernández-López, M., Gómez-Pérez, A., & Juristo, N. (1997). *Methontology: From ontological art towards ontological engineering*.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220. https://doi.org/10.1006/knac.1993.1008

Hor, A.-H., Jadidi, A., & Sohn, G. (2016). BIM-GIS INTEGRATED GEOSPATIAL INFORMATION MODEL USING SEMANTIC WEB AND RDF GRAPHS. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *III–4*, 73–79. https://doi.org/10.5194/isprsannals-III-4-73-2016

Hu, W., & Qu, Y. (2008). Falcon-AO: A practical ontology matching system. *Journal of Web Semantics*, *6*(3), 237–239.

Hu, Z.-Z., Leng, S., Lin, J.-R., Li, S.-W., & Xiao, Y.-Q. (2022). Knowledge Extraction and Discovery Based on BIM: A Critical Review and Future Directions. *Archives of Computational Methods in Engineering*, *29*(1), 335–356. https://doi.org/10.1007/s11831-021-09576-9

Hu, Z.-Z., Yuan, S., Benghi, C., Zhang, J.-P., Zhang, X.-Y., Li, D., & Kassem, M. (2019). Geometric optimization of building information models in MEP projects: Algorithms and techniques for improving storage, transmission and display. *Automation in Construction*, *107*, 102941. https://doi.org/10.1016/j.autcon.2019.102941

Huang, Z., & Zhao, W. (2020). Combination of ELMo Representation and CNN Approaches to Enhance Service Discovery. *IEEE ACCESS*, *8*, 130782–130796.

Hyun, K.-C., Min, S., Choi, H., Park, J., & Lee, I.-M. (2015). Risk analysis using fault-tree analysis (FTA) and analytic hierarchy process (AHP) applicable to shield TBM tunnels. *Tunnelling and Underground Space Technology*, *49*, 121–129. https://doi.org/10.1016/j.tust.2015.04.007

Jean-Mary, Y. R., Shironoshita, E. P., & Kabuka, M. R. (2009). Ontology matching with semantic verification. *Journal of Web Semantics*, *7*(3), 235–251.

Jiang, L., Shi, J., Wang, C., & Pan, Z. (2023). Intelligent control of building fire protection system using digital twins and semantic web technologies. *Automation in Construction*, *147*, 104728. https://doi.org/10.1016/j.autcon.2022.104728

Kampmann, J., Eskesen, S., & Summers, J. (1998). Risk assessment helps select the contractor for the Copenhagen Metro System. *Proceedings of the World Tunnel Congress*, *98*, 123–128.

Ke, S., Wei, W., & Shoubing, Z. (2015). Fuzzy analytical hierarchy process comprehensive evaluation method for health monitoring data of shield tunnels. *Journal of Disaster Prevention and Mitigation Engineering*, *35*(6), 769–776.

Khademi Hamidi, J., Shahriar, K., Rezai, B., Rostami, J., & Bejari, H. (2010). Risk assessment based selection of rock TBM for adverse geological conditions using Fuzzy-AHP. *Bulletin of Engineering Geology and the Environment*, *69*, 523–532.

Kiu, C.-C., & Lee, C.-S. (2006). Ontology mapping and merging through OntoDNA for learning object reusability. *Journal of Educational Technology & Society*, *9*(3), 27–42.

Kügler, P., Dworschak, F., Schleich, B., & Wartzack, S. (2023). The evolution of knowledge-based engineering from a design research perspective: Literature review 2012–2021. *Advanced Engineering Informatics*, *55*, 101892. https://doi.org/10.1016/j.aei.2023.101892

Le, T., & David Jeong, H. (2016). Interlinking life-cycle data spaces to support decision making in highway asset management. *Automation in Construction*, *64*, 54–64. https://doi.org/10.1016/j.autcon.2015.12.016

Leng, S., Lin, J.-R., Hu, Z.-Z., & Shen, X. (2020). A Hybrid Data Mining Method for Tunnel Engineering Based on Real-Time Monitoring Data From Tunnel Boring Machines. *IEEE Access*, *8*, 90430–90449. IEEE Access. https://doi.org/10.1109/ACCESS.2020.2994115

Li, J., Tang, J., Li, Y., & Luo, Q. (2009). RiMOM: A Dynamic Multistrategy Ontology Alignment Framework. *IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING*, *21*(8), 1218–1232. https://doi.org/10.1109/TKDE.2008.202

Li, L., Gu, P., & Wang, W. (2021). Ontology Modeling of Army Equipment Maintenance Support Business Process under the New System Base on IDEF5/Owl. *2021 IEEE International Conference on Power Electronics, Computer Applications (ICPECA)*, 331–335.

Li, T. H. Y., Ng, S. T., & Skitmore, M. (2013). Evaluating stakeholder satisfaction during public participation in major infrastructure and construction projects: A fuzzy approach. *Automation in Construction*, *29*, 123–135. https://doi.org/10.1016/j.autcon.2012.09.007

Lin, S.-S., Zhang, N., Zhou, A., & Shen, S.-L. (2022). Risk evaluation of excavation based on fuzzy decision-making model. *Automation in Construction*, *136*, 104143.

Ma, Z., Choi, J., Liu, P., & Sohn, H. (2022). Structural displacement estimation by fusing vision camera and accelerometer using hybrid computer vision algorithm and adaptive multi-rate Kalman filter. *Automation in Construction*, *140*, 104338. https://doi.org/10.1016/j.autcon.2022.104338

Mignard, C., & Nicolle, C. (2014). Merging BIM and GIS using ontologies application to urban facility management in ACTIVe3D. *Computers in Industry*, *65*(9), 1276–1290.

Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology*.

Nuñez, D. L., & Borsato, M. (2018). OntoProg: An ontology-based model for implementing Prognostics Health Management in mechanical machines. *Advanced Engineering Informatics*, *38*, 746–759. https://doi.org/10.1016/j.aei.2018.10.006

Pan, Y., Zhang, L., Wu, X., & Skibniewski, M. J. (2020). Multi-classifier information fusion in risk analysis. *Information Fusion*, *60*, 121–136. https://doi.org/10.1016/j.inffus.2020.02.003

Pardo, C., Pino, F. J., García, F., Piattini, M., & Baldassarre, M. T. (2012). An ontology for the harmonization of multiple standards and models. *Computer Standards & Interfaces*, *34*(1), 48–59. https://doi.org/10.1016/j.csi.2011.05.005

Pauwels, P., & Terkaj, W. (2016). EXPRESS to OWL for construction industry: Towards a recommendable and usable ifcOWL ontology. *Automation in Construction*, *63*, 100–133. https://doi.org/10.1016/j.autcon.2015.12.003

Pauwels, P., Zhang, S., & Lee, Y.-C. (2017). Semantic web technologies in AEC industry: A literature overview. *Automation in Construction*, *73*, 145–165. https://doi.org/10.1016/j.autcon.2016.10.003

Rasmussen, M. H., Pauwels, P., Lefrançois, M., Schneider, G. F., Hviid, C. A., & Karlshøj, J. (2017). *Building Topology Ontology*. http://rgdoi.net/10.13140/RG.2.2.32365.28647

Ren, G.-F., Chen, S., Zhou, C.-J., & Li, X.-D. (2023). Research on Health and Safety Assessment Method of Shield Tunnel Structure During Operation Period. *Journal of WUT(Information & Management Engineering)*, *45*, 520–526.

Shi, J., Pan, Z., Jiang, L., & Zhai, X. (2023). An ontology-based methodology to establish city information model of digital twin city by merging BIM, GIS and IoT. *Advanced Engineering Informatics*, *57*, 102114. https://doi.org/10.1016/j.aei.2023.102114

Venugopal, M., Eastman, C. M., & Teizer, J. (2015). An ontology-based analysis of the industry foundation class schema for building information model exchanges. *Advanced Engineering Informatics*, *29*(4), 940–957. https://doi.org/10.1016/j.aei.2015.09.006

Wang, Y., Liu, Y., Sun, Z., & Tang, P. (2018). *A Bayesian-entropy Network for Information Fusion and Reliability Assessment of National Airspace Systems*.

Yoon, S., & Yu, Y. (2017). Extended virtual in-situ calibration method in building systems using Bayesian inference. *Automation in Construction*, *73*, 20–30. https://doi.org/10.1016/j.autcon.2016.10.008

Zhang, W., Sun, K., Lei, C., Zhang, Y., Li, H., & Spencer, B. F. (2014). Fuzzy Analytic Hierarchy Process

Synthetic Evaluation Models for the Health Monitoring of Shield Tunnels. *Computer-Aided Civil and Infrastructure Engineering*, *29*(9), 676–688. https://doi.org/10.1111/mice.12091

Zhao, H., Yan, X., & Wang, G. (2019). Fault diagnosis of wind turbine generator based on deep autoencoder network and XGBoost. *Automation of Electric Power Systems*, *43*(1), 81–90.

Zhong, B., Gan, C., Luo, H., & Xing, X. (2018). Ontology-based framework for building environmental monitoring and compliance checking under BIM environment. *Building and Environment*, *141*, 127–142. https://doi.org/10.1016/j.buildenv.2018.05.046