IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 October 2023, accepted 15 December 2023, date of publication 18 December 2023,
date of current version 27 December 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3344320

RESEARCH ARTICLE

Machine Learning-Driven Ontological Knowledge Base for Bridge Corrosion Evaluation

YALI JIANG[1], HAIJIANG LI[2], GANG YANG[3], CHEN ZHANG[4], AND KAI ZHAO[5]
[1]College of Transportation Engineering, Dalian Maritime University, Dalian 116026, China
[2]School of Engineering, Cardiff University, CF24 3AA Cardiff, U.K.
[3]College of Transportation Engineering, Dalian Maritime University, Dalian 116026, China
[4]ZJYY (Dalian) Bridge Underwater Inspection Company Ltd., Dalian 116026, China
[5]School of Engineering and Architecture, University College Cork, Cork, T12 K8AF Ireland

Corresponding authors: Haijiang Li (lih@cardiff.ac.uk) and Chen Zhang (zhangyu_zc@163.com)

This work was supported in part by the China Scholarship Council under Grant CSC 202006570025, and in part by the Building Information Modeling (BIM) for Smart Engineering Centre in Cardiff University, U.K.

ABSTRACT In bridge maintenance, assessing structural performance requires adherence to rules outlined in safety and regulatory standards which can be effectively and formally represented in both human and machine-readable formats using ontologies. However, ontology-based semantic inference alone falls short when faced with the complicated mathematical operations required for structural analysis. The increasing digitization of bridge engineering has opened doors to data-driven prediction methods. Machine learning (ML)-based models, in particular, have the capacity to learn from historical data and forecast future structural performance with remarkable accuracy. This paper introduces an innovative approach that integrates ML models with an ontological knowledge base for evaluating bridge corrosion. Web Ontology Language and Semantic Web Rule Language are combined to develop the knowledge base. Random forest algorithm is used to train the ML model with a good agreement (coefficient of determination of 0.989 and root mean square error of 1.200). A Python-based module is designed to seamlessly integrate ML predictions with ontology-based semantic inference. The proposed approach not only infers the corrosion ratings based on the rules defined in the Network Rail standard, but also infers the structural safety performance based on predicted structural response under the action of corrosion. To demonstrate the effectiveness of the developed method in enabling accurate and rational evaluations, a real bridge in the UK is showcased as a practical application.

INDEX TERMS Knowledge engineering, knowledge base, machine learning, ontology, corrosion evaluation, bridge maintenance, data-driven, random forest.

I. INTRODUCTION

Bridges are a vital part of the architecture, engineering, and construction (AEC) industry, and effective maintenance is essential for ensuring good condition of their structures [1], [2]. Bridge maintenance tasks carry a profound responsibility, and they must strictly adhere to a complex web of safety and regulatory standards. Bridge maintenance standards are typically represented in a manner recognized by humans but then converted to a different format for storage in computers [3]. Although that format is computer-readable, computers cannot understand the content of the documents, which results in inefficiencies when the documents are used. Therefore, using the Semantic Web (SW) to facilitate the use of domain-specific knowledge has become the focus of intensive investigation.

The SW is a group of languages or technologies, such as Resource Description Framework (RDF) and Web Ontology Language (OWL), that allow machines to understand the meaning or semantics information on the World Wide Web [4]. OWL extends the capabilities of RDF by adding rich semantic expressions, and it is

The associate editor coordinating the review of this manuscript and approving it for publication was Weiping Ding.

© 2023 The Authors. This work is licensed under a Creative Commons Attribution 4.0 License.
For more information, see https://creativecommons.org/licenses/by/4.0/

a semantic language for developing and sharing ontologies. The core ontologies are knowledge representations of a domain that contain explicit description of concepts, attributes or features of those concepts, and logical restrictions on them. The objective of an ontology is to represent knowledge in a specific domain in a both human and machine-readable format. Distinct from analytics and algorithms, there are several advantages of ontology, including: facilitating knowledge sharing and reuse, supporting interoperability, enabling automated inference, improving information retrieval, and enhancing consistency and accuracy [5], [6].

Existing ontologies related to bridge maintenance tasks can effectively integrate static information from industry manuals and norms. For instance, Li et al. [7] presented a bridge structure and health monitoring knowledge base that enabled fine-grained modeling and domain-knowledge discovery. Moreover, Ren et al. [6] grouped semantically related bridge components and subsequently embedded rules to evaluate conditions for bridge maintenance operations. Similarly, Liu and EL-Gohary [8] used the SW and natural language processing (NLP) to solve problems related to bridge deterioration and inspections. However, the inadequacy of ontology-based semantic inference alone in managing the complex mathematical operations required for structural analysis. Limited research exists on dynamically linking semantic inference to information from data-driven methods.

Embracing advanced information and communications technologies has continually enhanced the intelligence level of bridge maintenance. The increasing availability of maintenance data from various sources enables numerous data-driven prediction methods. Recent research studies [9], [10], [11] have emphasized the importance of data-driven bridge performance prediction in supporting bridge maintenance decision-making. Machine learning (ML)-based models, in particular, have the capacity to learn from historical data and forecast future structural performance with high accuracy. For example, Jaafaru and Agbelie [12] introduced a bridge maintenance planning framework (BMPF) that combines random forest (RF) algorithm, multi-attribute utility theory, and genetic algorithms to help engineers evaluate and maintain bridges effectively. The study analyzed 95 bridges, achieving an 84% accuracy in ML predictions. Liu et al. [13] developed an artificial neural network (ANN)-based method for rapid seismic fragility assessment of regular bridges with root mean absolute error of 0.173 and coefficient of determination of 0.997. The study indicated that the ML method is an effective alternative for seismic assessment of bridges with significantly reduced computation time. Therefore, this paper introduces an approach that combines ML techniques with an ontological knowledge base for the evaluation of corrosion severity and extent in bridges. The fusion of different technologies, such as this, can significantly enhance assessment processes and optimize bridge maintenance.

II. LITERATURE REVIEW

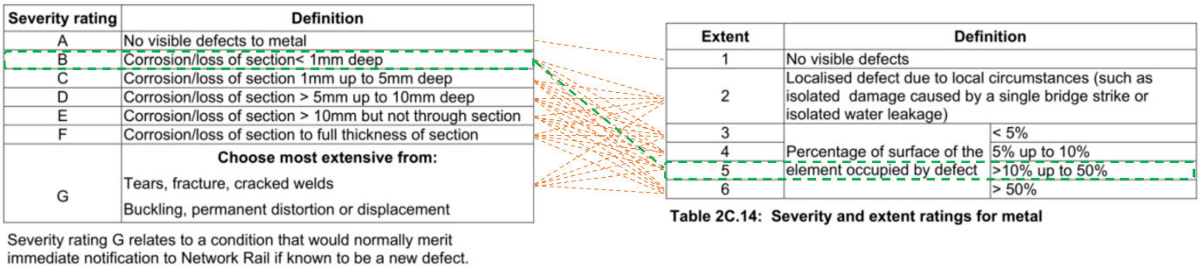
A. CORROSION EVALUATION STANDARDS

In the context of railway infrastructure, regular inspection and assessment of bridges are essential to ensure their structural integrity and overall condition. Corrosion on railway bridges represents a significant concern, posing a direct threat to their safety and structural integrity. This corrosion primarily arises from the interaction between metal components, typically steel, and environmental factors such as moisture, oxygen, pollutants, and temperature fluctuations, initiating oxidation and material degradation. It can manifest in various forms, including surface corrosion, pitting corrosion, uniform corrosion, galvanic corrosion, and stress corrosion cracking, all of which can weaken the metal components and reduce their load-bearing capacity. If not addressed promptly, corrosion can lead to structural failures.

To address this issue, railway authorities and maintenance teams in the UK employ various strategies, with bridge corrosion evaluation being a crucial aspect. This process involves assessing the extent and severity of corrosion that may affect various structural elements of a bridge, such as beams, columns, and support structures. Figure 1 illustrates the method outlined in the Network Rail standard (NR/L3/CIV/006/2C). The severity rating is determined based on the depth of corrosion, while the extent rating is calculated by considering the percentage of the element's surface area occupied by corrosion. Severity assessment helps in understanding the level of structural degradation and potential safety risks, while extent measurement is essential for determining the overall impact on the bridge's structural integrity. For instance, if a corrosion depth is 0.9mm and the corrosion occupies 20% of the element's surface, it will be categorized as severity B and extent 5 according to this standard.

It is worth noting that severity ratings of A, F and G can directly determine the structural condition. Severity A indicates the absence of visible corrosion defects. Severity F indicates the presence of severe corrosion defects that have impacted normal functions. Severity G necessitates immediate notification to Network Rail. For others, they can provide a rough quantitative representation of defect ratings, but they cannot accurately determine whether defects have a substantial impact on structural safety performance. Maintenance decisions often rely heavily on the subjective judgment of engineers.

Finite element models enable a detailed analysis of structural behavior under various conditions [14]. Using the depth and area of corrosion as inputs, a structural analysis is conducted through the finite element method to assess the potential consequences and likelihood of failure or damage. Data from the finite element analysis is then used to develop a real-time ML surrogate model for accurately predicting the structure's response. This analysis informs decision-making regarding maintenance, repair, or replacement.



Severity rating	Definition
A	No visible defects to metal
B	Corrosion/loss of section< 1mm deep
C	Corrosion/loss of section 1mm up to 5mm deep
D	Corrosion/loss of section > 5mm up to 10mm deep
E	Corrosion/loss of section > 10mm but not through section
F	Corrosion/loss of section to full thickness of section
G	**Choose most extensive from:** Tears, fracture, cracked welds; Buckling, permanent distortion or displacement

Severity rating G relates to a condition that would normally merit immediate notification to Network Rail if known to be a new defect.

Extent	Definition	
1	No visible defects	
2	Localised defect due to local circumstances (such as isolated damage caused by a single bridge strike or isolated water leakage)	
3	Percentage of surface of the element occupied by defect	< 5%
4		5% up to 10%
5		>10% up to 50%
6		> 50%

Table 2C.14: Severity and extent ratings for metal

FIGURE 1. Severity and extent ratings for metallic elements of bridges.

The loading methods to be employed in the analysis process align with the Network Rail standard (NR/GN/CIV/025). The rail traffic live load is determined based on its route availability (RA) number. RA numbers generally range from the lowest capacity RA0 to the highest at RA15 represented by 25 British Standard Units (BSUs) of Type RA1 loading. The static loading for 20 units of Type RA1 loading is depicted in Figure 2. In the case of simply supported spans, type RA1 loading can be expressed as an equivalent uniformly distributed load (EUDL).

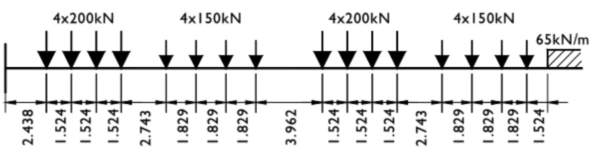


FIGURE 2. 20 units of type RA1 loading.

To account for dynamic effects, a factor of $(\varphi_1 + \varphi_{11})$ is applied to EUDL loading to allow for impact, oscillation and other dynamic effects, including those caused by track and wheel irregularities, given that the track is designed for speeds less than 100 mph. For the end shear, a factor of 2/3 $(\varphi_1 + \varphi_{11})$ is applied. φ_1 and φ_{11} can be calculated refer to (1) – (3).

$$\varphi_1 = \frac{k}{1 - k + k^4} \tag{1}$$

where

$$k = \frac{v}{4.47 L_\phi n_0} \tag{2}$$

where v is the permissible speed on the bridge. L_ϕ is the determinant length of centre to centre of supports in metres. n_0 is the fundamental natural frequency of the structure.

And for φ_1

$$\varphi_{11} = \alpha \left[56 e^{-\left(\frac{L_\phi}{10}\right)^2} + 50 \left(\frac{L n_0}{80} - 1 \right) e^{-\left(\frac{L_\phi}{10}\right)^2} \right] \tag{3}$$

where L is the span of the member, centre to centre of supports in metres. $\alpha = 0.0002v$.

B. ONTOLOGIES IN THE BRIDGE DOMAIN

Ontology is a new semantic technology being widely adopted in various domains, including knowledge engineering, natural language processing, collaborative information systems, intelligent information integration, internet intelligent information acquisition, and knowledge management [15], [16], [17]. Studies related to bridge ontology models can be categorized into two distinct groups: 1) ontology-based knowledge management and information retrieval; and 2) logical inference for holistic decision-making.

The first group focuses on developing domain-specific ontologies to manage and organize knowledge, making it easier to retrieve relevant information, and share domain-specific knowledge. For example, Wu et al. [5] developed an ontological knowledge base called CBRPMO for concrete bridge rehabilitation project management, aiming to enhance information integration and constraint management. Liu et al. [8], [18] introduced an ontology named BridgeOnto to represent information related to bridge inspections. Li et al. [7] proposed the BSHM (Bridge Structural Health Monitoring) ontology for bridge health monitoring systems, with the aim of supporting the integration of heterogeneous sensor data and the discovery of domain knowledge. Zhang and El-Diraby [19] created the AR-Onto, focusing on the AEC industry, including bridges, to tackle challenges related to information exchange and knowledge sharing. Compton et al. [20] presented an OWL2 ontology named SSN, designed to describe sensors and observations concerning capabilities, measurement processes, observations, and deployments. Banujan and Vasanthapriyan [21] created a bridge maintenance ontology aimed at facilitating the sharing of bridge maintenance knowledge. Ontologies play a central role in structuring data and facilitating efficient knowledge management and retrieval in these studies.

Studies in the second group focus on employing ontology-based knowledge representation and semantics-to-reasoning techniques to improve efficiency, automation, and comprehensiveness in decision-making processes for engineers. For example, an expert system was developed by Becker and Gebbeken [22] to aid engineers in assessing the condition and planning maintenance for aging bridges, utilizing an ontology-based knowledge representation to store expert knowledge and inference algorithms, with a practical demonstration involving a reinforced concrete bridge assessment. Ren et al. [6] presented a holistic approach utilizing ontology-based knowledge representation to improve bridge maintenance efficiency during the operation stage, providing automatic rule checking and enhancing knowledge management, communication, and decision-making for engineers. This approach was validated through semantic validation, syntactical validation, and case study validation, underscoring its capability to integrate diverse domain knowledge and enable more comprehensive decision-making in bridge management. Zhang et al. [23] introduced a holistic design approach for energy pile bridge deicing systems, addressing the limitations of traditional single-domain and objective-oriented methods by incorporating factors like life-cycle cost, investment payback cycle, and carbon emissions. They presented "OntoBDDS", a tool based on ontology and Semantic Web Rule Language (SWRL) rules, which automates the provision of financial, safety, and heat flux information to aid designers in evaluating and optimizing deicing system designs during the early bridge design stage. Chai and Wang [24] developed an evaluation and decision-making framework for concrete surface quality that combines computer vision and ontology. By utilizing ontology and a defect identification quantification model, they used ontology reasoning technology to intelligently evaluate and make decisions regarding concrete surface quality, bridging the gap between low-level semantics acquired by computer vision and high-level semantics understood by humans from images.

Both groups search for information by using the SPARQL Protocol and RDF Query Language (SPARQL) designed and endorsed by the World Wide Web Consortium (W3C), which is an international community that develops open standards to ensure the long-term growth of the web. SWRL and Semantic Query-Enhanced Web Rule Language (SQWRL) were created as supplements to OWL to represent the reasoning rules in some research. SQWRL further embeds SPARQL queries in rules, thereby achieving reasoning and querying simultaneously. However, research on the implementation of semantic reasoning processes dynamically linking to information is inadequate, such that information from third-party applications cannot be invoked during inference.

III. DEVELOPMENT BRIDGE CORROSION EVALUATION ONTOLOGY (BCEO)

The research framework of ontology development includes ontology specification, knowledge acquisition, conceptualization, implementation, and evaluation [25]. The methodology used is a combination of two approaches: the "methontology" approach [26] and the "Uschold and Gruninger" ontology building approach [27]. The "Uschold and Gruninger" approach provides a thorough explanation of how to define the purpose and scope, formalize, evaluate, and document ontologies. The methontology approach is more detailed and focused on how to acquire, conceptualize, and implement knowledge. Together, these two approaches create a comprehensive methodology for building ontologies. Figure 3 outlines the different activities, keys and tasks involved in the process.

As illustrated in Figure 3, it is priority to initially define the ontology's scope and purpose. As emphasized by Uschold and Gruninger [27], the development of ontologies is often motivated by practical use cases. In this research, the motivating scenario revolves around the River Neath Swing Bridge, a railway bridge presently undergoing extensive maintenance due to severe corrosion affecting its structural elements. A survey has revealed that the majority of metallic railway bridges in the UK – approximately 10,000 in total – were constructed a century ago, facing similar challenges to the case bridge [28]. So, the application domain of the BCEO is bridge corrosion evaluation. The BCEO is designed to connect ML-based forecast results to provide accurate judgments of bridge safety.

Subsequently, knowledge capture and taxonomy of relevant terms were conducted. To avoid ambiguity and facilitate the later expansion of the BCEO, an analysis was conducted on standards NR/L3/CIV/006 and NR/GN/CIV/025 mentioned in the literature review. This analysis involved collecting unified terminology, encompassing the major and minor elements, and quantitative condition ratings to elements. Following this phase, the BCEO was formally coded using OWL in a semantic, computational logic-based format through Protégé. The coded BCEO then underwent logical validation and was subsequently implemented in a case study.

A UML (Unified Modelling Language) diagram in Figure 4 depicts the highest-level terms within the BCEO, comprising 14 core classes, 19 object properties, and 23 key data properties. Classes serve to systematically organize and categorize knowledge and can be further divided into subclasses, enabling a hierarchical knowledge structure. For instance, following the "whole to part" principle, the "Element" class is subdivided into "major element" and "minor element" classes. Properties, on the other hand, refer to attributes or relationships that define how classes or individuals are interconnected or characterized. They are a fundamental building block for creating well-structured and expressive ontologies that can support various applications, including data integration, reasoning, and knowledge discovery. For example, the object properties "buildBy" and "managedBy" connect individuals belonging to the class "Bridge" to individuals belonging to the class "Organization", resulting in the corresponding RDF triples: "Bridge, buildBy, Organization" and "Bridge, managedBy, Organization". Depending on the specific requirements of this research, properties with characteristic setting, quantifier restrictions, cardinality restriction, domain and range restriction are created to describe characteristics of various individuals in both a quantitative and qualitative way. For example:

- characteristic settings: cooperateWith is Symmetric

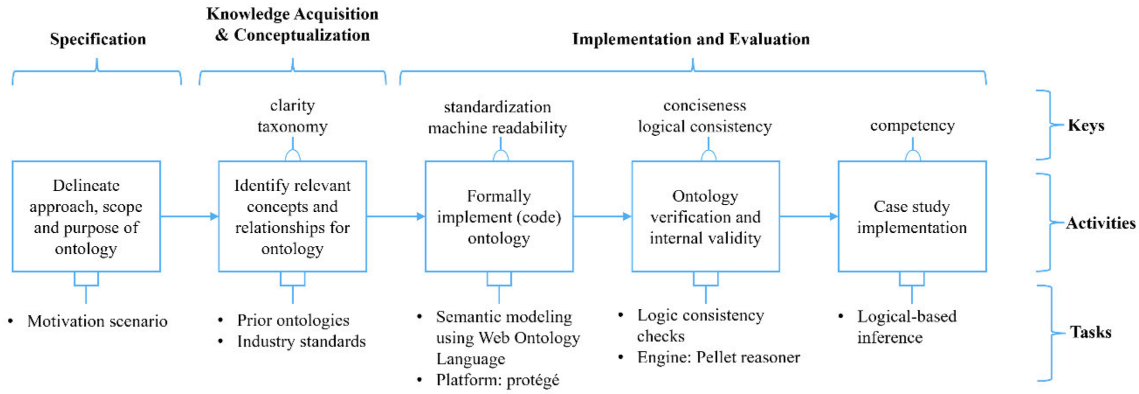


FIGURE 3. Research framework for ontology development.

- existential restrictions: hasElement some Element
- universal restrictions: hasElement only (Element or MiscellaneousItem)
- qualified cardinality restrictions: spanNumder exactly 1 xsd:int
- domain restrictions: purpose domain Standard
- range restrictions: address range xsd:string

The information described above is formally represented using the OWL formal language, and a URL (Uniform Resource Locator) in the form of https://w3id.org/BCEO has been incorporated. With this setup, the BCEO can employ built-in reasoners like Pallet to initiate logical inference and validation. In Figure 5, an example of semantic inference is illustrated. Initially, various individuals and their relationships are established (shown in blue). For instance, there is an individual named ''RiverNeathSwingBridge'' associated with a major element called ''Deck1''. Additionally, a main girder named ''LongitudinalMainGirder(exposed)1'' is assigned to ''Deck1''. Further properties such as ''spanNumber'', ''hasStakeholders'' and ''cooperateWith'' are also added. In line with the semantics defined earlier, when the reasoner is synchronized, implicit knowledge becomes explicitly inferred. For instance, utilizing the transitive characteristic of the ''hasElement'' property, the BCEO can infer that ''RiverNeathSwingBridge'' possesses an element known as ''LongitudinalMainGirder(exposed)1''. And any inconsistencies are promptly identified and highlighted in red, while consistent results are marked in yellow. Consistent information can be continuously added to the BCEO, further enhancing its depth and breadth.

In addition to making implicit information explicit, advanced deductive reasoning capabilities are essential for assessing corrosion severity and extent. SWRL provides a mechanism for expressing complex relationships and logical constraints that enhance ontology capabilities within the SW. It also involves mathematical operations and constraints, enabling more comprehensive and expressive knowledge representation based on mathematical relationships. As listed in Table 1, 12 SWRL rules were created to express the corrosion evaluation method defined in standards as illustrated in Figure 1.

An SWRL rule is composed of two primary parts: the antecedent (body), which is located on the left side, and the consequent (head), located on the right side. These two parts are connected by the symbol ''→''. SWRL offers several types of atoms, including class atoms, individual property atoms, data valued property atoms, and built-in atoms. By connecting atoms with ''∧''' and utilizing variables denoted by ''?'', the satisfaction of atoms in the antecedent leads to the truth of atoms in the consequent. Table 2 provides a list of several atoms used in this research. These atoms in SWRL rules can take forms such as C(x), P(x,y), sameAs(x,y), or differentFrom(x,y), where C is an OWL description, P is an OWL property, and x and y are either variables, OWL individuals or data values. Indeed, the ontology becomes undecidable when extended in this manner, as rules can be employed to emulate role value maps.

IV. DEVELOPMENT OF ML MODEL

The River Neath Swing Bridge, designated as a Grade II listed six-deck underbridge, carries two lines of the SDI1 railway over the River Neath. Among its six decks, five are simply supported, while the former swing span deck (mechanism no longer operational) was initially designed to pivot open on the central support. The swing span comprises a riveted steel bowstring through truss, cross girders, and a steel deck, with overhead lateral restraints at midspan. The substructure consists of wrought iron cylindrical braced piers with concrete infill, masonry abutments, and wingwalls. Each of the five simply supported spans comprises two longitudinal steel plate girders, cross girders, rail bearers, and a steel deck, supported on pairs of braced cylindrical piers or masonry abutments. The bridge spans vary in length, with 16.81 m (Span 1), 17.38 m (Span 2), 17.01 m (Span 3), 49.91 m (Span 4), 16.90 m (Span 5), and 13.08 m (Span 6).

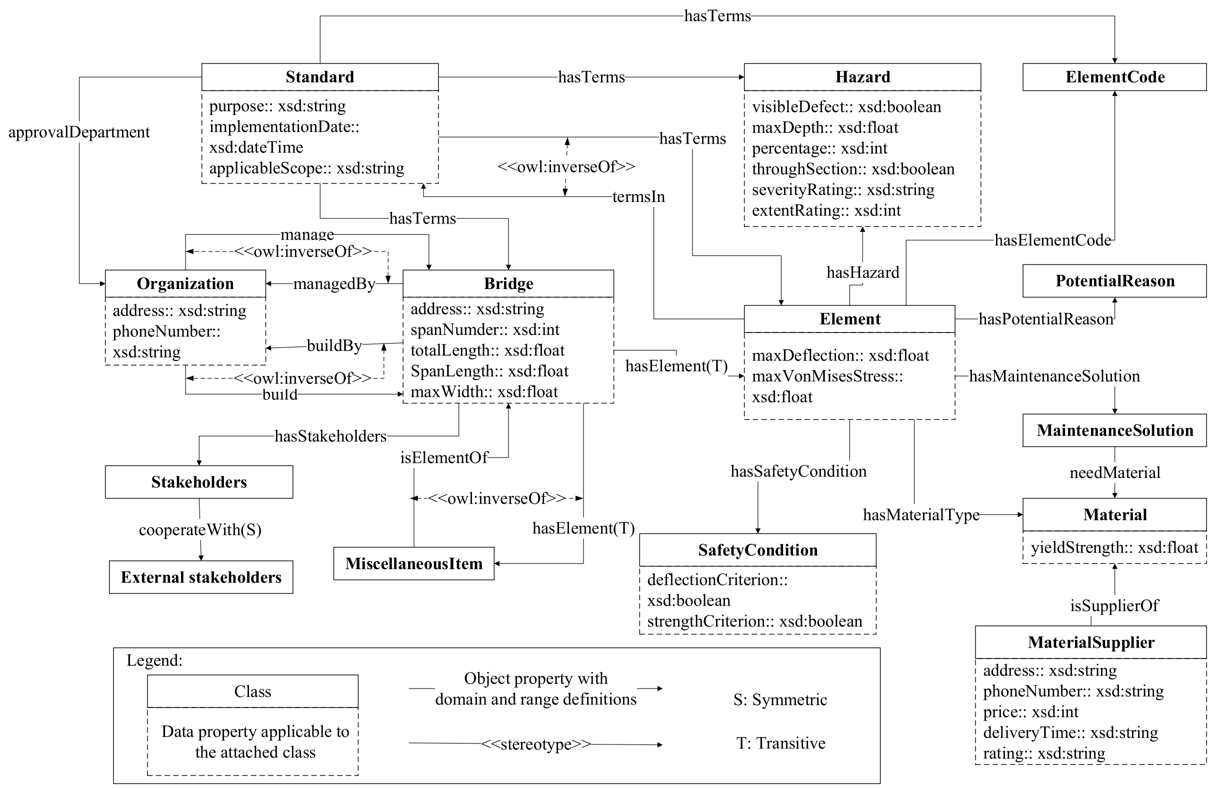


FIGURE 4. The high-level overview of the BCEO.

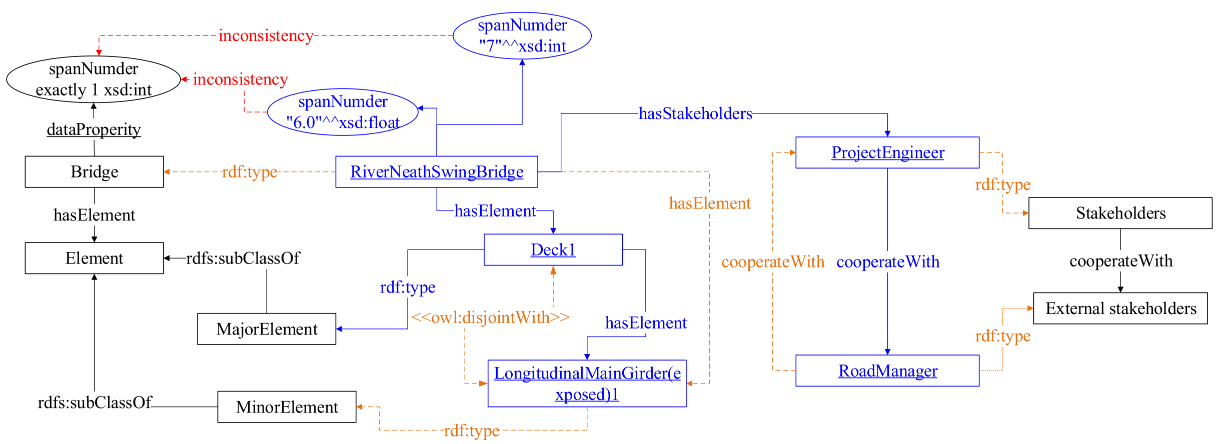


FIGURE 5. An example of semantic inference.

As depicted in Figure 6, the finite element model and the corresponding ML surrogate model are developed based on one of the longitudinal steel plate girders, a critical structural component of Span 2. The finite element model is built using a combination of 2D triangles and rectangular shell element (S3R and S4R) with five integration points with average edge size of 150 mm. Structural loading adheres to the guidelines specified in standard NR/GN/CIV/025. The RA number of the River Neath Swing Bridge is the specified rating of RA8 (at 20 mph) and heavy axle rating RA10 (at 20 mph). For RA10, the static loading for EUDL and end shear amounts to 162.0 kN and 94.7 kN for each track, respectively. The fundamental natural frequency of the structure is measured at 16.8 Hz using a vibration camera. In accordance with equations 1-3, the values of k, φ_1, and φ_{11} are 0.01532, 0.01556, and 0.03677, respectively. Thus, the total load applied to the tested member is calculated as 170.5 kN for EUDL and 98.0 kN for end shear.

Two parameters are designated as variables for the structure analysis: the corrosion depth and area. These selections

TABLE 1. SWRL rules for corrosion severity and extent evaluation.

Application 1: SWRL rules for determining the severity rating		
Rule1	Corrosion(?C)^visibleDefect(?C,false) -> severityRating(?C,"A")	If there are no visible defects, the severity rating is A.
Rule2	Corrosion(?C)^maxDepth(?C,?Cd)^swrlb:lessThan(?Cd,1)^throughSection(?C,false) -> severityRating(?C,"B")	If a corrosion is less than 1 mm deep, the severity rating is B.
Rule3	Corrosion(?C)^maxDepth(?C,?Cd)^swrlb:greaterThanOrEqual(?Cd,1)^swrlb:lessThanOrEqual(?Cd,5)^throughSection(?C,false) -> severityRating(?C,"C")	If a corrosion is 1mm up to 5mm deep, the severity rating is C.
Rule4	Corrosion(?C)^maxDepth(?C,?Cd)^swrlb:greaterThan(?Cd,5)^swrlb:lessThanOrEqual(?Cd,10)^throughSection(?C,false) -> severityRating(?C,"D")	If a corrosion is greater than 5mm up to 10mm deep, the severity rating is D.
Rule5	Corrosion(?C)^maxDepth(?C,?Cd)^swrlb:greaterThan(?Cd,10)^throughSection(?C,false) -> severityRating(?C,"E")	If a corrosion is greater than 10mm deep but not through section, the severity rating is E.
Rule6	Corrosion(?C)^throughSection(?C,true) -> severityRating(?C,"F")	If a corrosion reaches full thickness of section, the severity rating is F.
Application 2: SWRL rules for determining the extent rating		
Rule7	Hazard(?C)^percentage(?C,?Ca)^swrlb:equal(?Cd,0) -> extentRating(?C,1)	If the percentage of surface of the element occupied by corrosions is equal to 0%, the extent rating of risks is 1.
Rule8	Hazard(?C)^percentage(?C,?Ca)^swrlb:lessThan(?Ca,1) -> extentRating(?C,2)	If the percentage of surface of the element occupied by corrosions is less than 1%, the extent rating of risks is 2.
Rule9	Hazard(?C)^percentage(?C,?Ca)^swrlb:greaterThanOrEqual(?Ca,1)^swrlb:lessThan(?Ca,5) -> extentRating(?C,3)	If the percentage of surface of the element occupied by corrosions is no less than 1% and less than 5%, the extent rating of risks is 3.
Rule10	Hazard(?C)^percentage(?C,?Ca)^swrlb:greaterThanOrEqual(?Ca,5)^swrlb:lessThanOrEqual(?Ca,10)-> extentRating(?C,4)	If the percentage of surface of the element occupied by corrosions is 5% up to 10%, the extent rating of risks is 4.
Rule11	Hazard(?C)^percentage(?C,?Ca)^swrlb:greaterThan(?Ca,10)^swrlb:lessThanOrEqual(?Ca,50) -> extentRating(?C,5)	If the percentage of surface of the element occupied by corrosions is greater than 10% and no greater than 50%, the extent rating of risks is 5.
Rule12	Hazard(?C)^percentage(?C,?Ca)^swrlb:greaterThan(?Ca,50) -> extentRating(?C,6)	If the percentage of surface of the element occupied by corrosions is greater than 50%, the extent rating of risks is 6.

align with the standard NR/L3/CIV/006 for consistency. Specifically, the depth range for severity rating spans from 0 mm to 10 mm, encompassing ratings from B to E. Additionally, the extent rating, which is based on the percentage of surface occupied by corrosion, spans from 0% to 100%, covering ratings from 3 to 6. Given that the standard NR/L3/CIV/006 does not specify corrosion locations,

TABLE 2. Examples of atoms used in this research.

Atom type	Atom	Corresponding OWL element
Class atom	Hazard (?C)	Hazard (class)
Data valued property atom	severityRating (?C,?Cs)	severityRating (data-type property)
	extentRating (?C,?Ce)	extentRating (data-type property)
Individual property atom	hasMaterialType (?B,?BM)	hasMaterialType (object property)
Built-in atom	swrlb:greaterThanOrEqual (?Cd,1)	
	swrlb:lessThan(?Cd,5)	

all corrosions are concentrated in the center of the bridge component, as this is presumed to induce the most significant deformation and stress. The output of the model includes the maximum deflection and the maximum Von Mises stress in the girder considering structural deflection criterion and strength criterion, respectively. In this research, 1120 sets of simulation data are generated and utilized for training a ML surrogate model.

The ML model is trained using the RF method which is widely appreciated for its high accuracy, robustness, and ease of use in bridge engineering [29]. The RF model operates through an ensemble learning strategy, wherein it constructs multiple decision trees (DT) and combines their predictions to enhance accuracy [30]. Each DT utilizes a tree-like structure, with internal nodes representing attributes and branches indicating potential attribute values [31]. Decision rules are derived from these attributes to make predictions. The core steps in executing the RF model are illustrated in Figure. 7.

The process is as follows:

1) Assume that the original data set consists of N sets of data with M-dimensional features in each. Using the bootstrap sampling method, K different sample datasets are randomly selected from the original dataset as the sub-training set for each decision tree.
2) From the sampled subsets, m (where $0 < m < M$) features are chosen randomly to form the split feature set for each decision tree node. The optimal features within this subset are identified based on minimizing the Gini index, ensuring precise node splitting and branching.
3) Decision trees grow recursively from the root to leaf nodes until reaching a minimum leaf node size, with this process replicated across all trees in the RF ensemble.
4) Test data are input into the RF model, which generates separate predictions using K decision trees. The final prediction result is computed as the average of these individual tree predictions.

For the development of the model, the python scikit-learn package which offers a variety of ML functions is employed. 90% of the data randomly selected from 1120 sets of

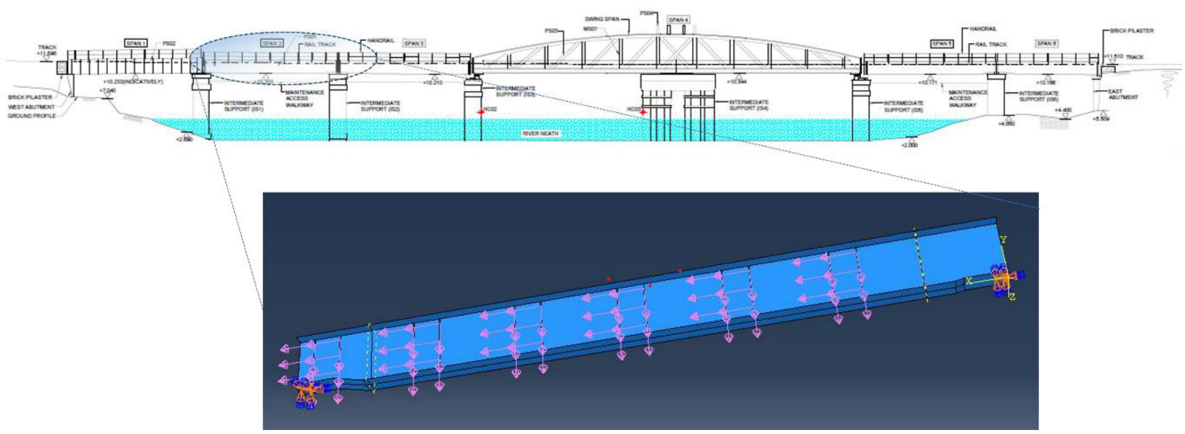


FIGURE 6. The structural analysis model for span 2 of the River Neath Swing Bridge in ABAQUS.

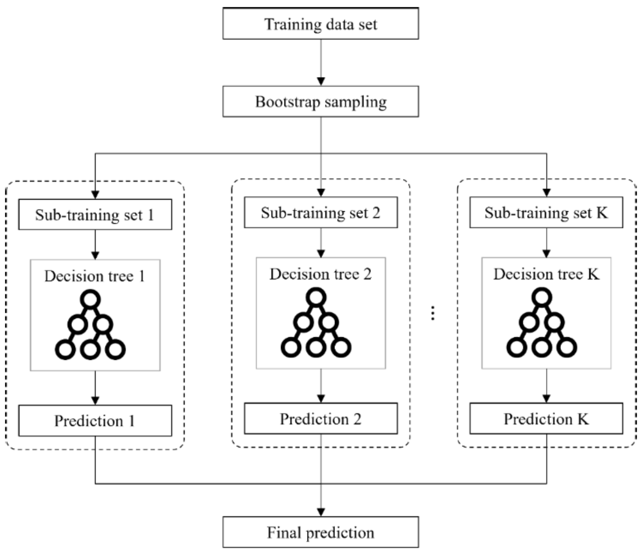


FIGURE 7. Core steps of the RF model execution.

simulation data is used as the training set of the RF, and the remaining 10% is used as the test set of the model. Key hyperparameters for training are determined based on trial-and-error procedure. When the number of trees in the forest is set to 50 and the maximum tree depth is configured as 14, the model yields satisfactory results.

The model is assessed using two evaluation metrics: R^2 (coefficient of determination) and RMSE (root mean square error). R^2 is a statistical measure used to assess the goodness of fit of a regression model. It quantifies the proportion of the variance in the dependent variable that can be explained by the independent variables in the model. The R^2 value typically ranges from 0 to 1, where $R^2 = 0$ means that the model fails to explain any variance in the dependent variable, while $R^2 = 1$ indicates a perfect fit, with the model accounting for all the variances. For R^2 values between 0 and 1, the model explains some, but not all, of the variance in the data, and higher R^2 values denote a superior fit. On the other hand, RMSE is a metric utilized to measure the average magnitude of errors between predicted values and the actual or observed values in a dataset. RMSE yields a single numerical value, shedding light on how closely the model's predictions align with the observed data. Smaller RMSE values indicate that the model's predictions closely match the observed values, implying a stronger fit of the model to the data. Conversely, larger RMSE values indicate that the model's predictions exhibit more significant errors and are less accurate. The detailed results of this evaluation are visually presented in Figure 8.

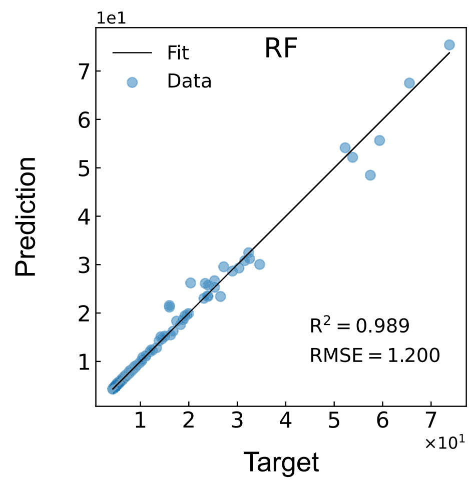


FIGURE 8. Results of R^2 and RMSE for RF.

RF achieves an R^2 of 98.9%, signifying a close correspondence between its predictions and the solutions obtained from finite element analysis. Additionally, it exhibits a low RMSE of 1.200, indicating its exceptional precision. This suggests that RF can offer highly accurate predictions, even when dealing with data that displays significant variability. Consequently, the model proves to be a reliable choice for making predictions in this scenario.

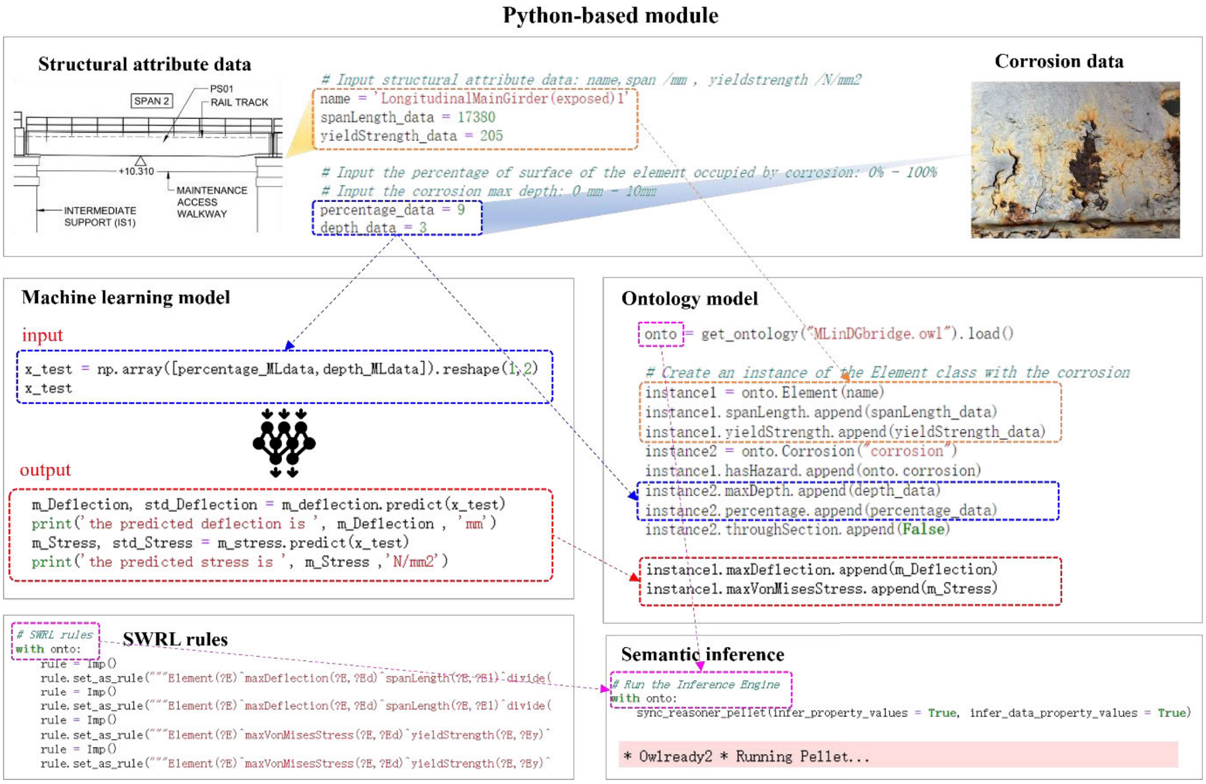


FIGURE 9. Running mechanism of the python-based module.

```
* Owlready * Adding relation C:\Users\19638\pythonCode\ML\MLinDGbridge.LongitudinalMainGirder(exposed)1 deflectionCriterion true
* Owlready * Adding relation C:\Users\19638\pythonCode\ML\MLinDGbridge.LongitudinalMainGirder(exposed)1 strengthCriterion true
* Owlready * Adding relation C:\Users\19638\pythonCode\ML\MLinDGbridge.corrosion extentRating 4.0
* Owlready * Adding relation C:\Users\19638\pythonCode\ML\MLinDGbridge.corrosion severityRating C

* Owlready2 * Pellet took 1.1111209392547607 seconds

The severity rating of corrosion on LongitudinalMainGirder(exposed)1 is ['C'] and the extent rating is [4.0].
For safety condition, its max deflection meets the requirements of deflection criterion:[True]; its max Von Mises Stress meets the
requirements of strength criterion: [True].
```

FIGURE 10. Reasoning results of corrosion evolution.

V. INTEGRATION ML MODEL WITH LOGICAL INFERENCE

A Python-based module is designed to seamlessly integrate ML predictions with ontology-based semantic inference, thereby facilitating the automatic evaluation of corrosion severity and extent. This module encompasses the trained ML model, which has the capability to provide real-time predictions of structural responses under the influence of actual corrosion. Furthermore, this module adopts an ontology-oriented programming approach, enabling the loading of OWL ontologies as Python objects. It also allows for modifications to these ontologies and supports automatic reasoning based on the ML prediction results.

Figure 9 provides a visual representation of the module's operational mechanism. Initially, it takes structural attribute data and corrosion data as input, and these data are then mapped to the relevant classes and attributes within the BCEO. This mapping process entails identifying the ontology classes and attributes that correspond to the input data and establishing associations with RDF triples. For example, the code "instance1 = onto.Element(name)" is used to associate a specific name with a class within an ontology. In this code:

- "onto" refers to an object representing a predefined BCEO.
- "Element" refers to a class defined within the BCEO.
- "name" refers to a variable employed to indicate the name for the instance that is being generated.

This code generates an instance "LongitudinalMainGirder(exposed)1" of that class "Element" and assigns it to the variable "instance1".

Additionally, the corrosion data serves as input for the ML model, enabling predictions of structural responses in the occurrence of corrosion, which encompasses estimating the maximum deflection and the maximum Von Mises stress.

Following these predictions, the resulting ML outcomes are mapped into the BCEO using the same method mentioned earlier, as demonstrated by the following code:

- instance1.maxDeflection.append(m_Deflection)
- instance1.maxVonMisesStress.append(m_Stress)

After the mapping process, additional SWRL rules are established to determine that the structural response levels adhere to acceptable limits based on relevant standards. As listed in Table 3, these rules assess two critical aspects: deflection and strength. Regarding deflection, the rule specifies that the maximum deflection should be less than 1/400 of the span. This criterion is based on information defined in the standard GB 50017-2017. Regarding strength, the rule specifies that the maximum Von Mises stress should be less than or equal to the yield strength of the structural material.

TABLE 3. SWRL rules for reasoning the safety condition of the structure.

Application 3: SWRL rules for the deflection criterion	
Rule13	Element(?E)^maxDeflection(?E,?Ed)^spanLength(?E,?El)^swrlb:divide(?div,?El,400)^swrlb:lessThan(?Ed,?div)-> deflectionCriterion(?E,true)
Rule14	Element(?E)^maxDeflection(?E,?Ed)^spanLength(?E,?El)^swrlb:divide(?div,?El,400)^swrlb:greaterThanOrEqual(?Ed, ?div)-> deflectionCriterion(?E,false)
Application 4: SWRL rules for the strength criterion	
Rule15	Element(?E)^maxVonMisesStress(?E,?Ed)^yieldStrength(?E,?Ey)^swrlb:lessThan(?Ed,?Ey)-> strengthCriterion(?E,true)
Rule16	Element(?E)^maxVonMisesStress(?E,?Ed)^yieldStrength(?E,?Ey)^swrlb:greaterThanOrEqual(?Ed,?Ey)-> strengthCriterion(?E,false)

Finally, deductive reasoning is performed based on all SWRL rules by running the inference engine. Pellet is used as the reasoner for this research. Given the structural attribute data and corrosion data as input (Figure 9), corrosion severity and extent ratings, and the safety condition, are inferred automatically and shown in Fig. 10. The corrosion severity rating for "LongitudinalMainGirder(exposed)1" is C, and the extent rating is 4.0. When assessing the impact of this corrosion, it is displayed that its maximum deflection aligns with the deflection criterion, and its maximum Von Mises Stress satisfies the strength criterion.

Overall, this integrated approach leverages Python libraries such as Owlready2 and Scikit-Learn to enhance the module's functionality and capabilities. Each of these libraries plays a distinct role in the overall process, with Owlready2 facilitating ontology operations, Scikit-Learn supporting ML tasks. This integration, on one hand, significantly enriches the existing static knowledge base by dynamically linking to real-time ML information, with a primary focus on ensuring bridge structural safety as the paramount consideration, thereby enabling more precise and accurate decision-making. On the other hand, ontology technology lends its support to data-driven predictions by furnishing domain-specific knowledge, ultimately facilitating the prediction of outcomes that meet stringent safety and regulatory requirements.

VI. CONCLUSION

This study has presented an innovative approach that effectively combines ML techniques with ontology-based knowledge representation to enhance assessing structural performance in the field of bridge maintenance, with a specific focus on corrosion severity and extent evaluation. The first contribution of this paper is the development of a bridge corrosion evaluation ontology (BCEO), providing a systematic knowledge base for managing various factors and attributes related to bridge corrosion defects. The second contribution facilitates improving the accuracy and precision of decision-making related to bridge structural safety, by dynamically linking real-time ML information with a static knowledge base. The use of ontology technology has provided domain-specific knowledge that supports data-driven predictions, ensuring compliance with stringent safety and regulatory requirements.

The practical application on a real bridge in the UK showcased the practicality and effectiveness of the developed method. The results indicate that this approach has the capacity to offer more precise and rational decision support in bridge maintenance. However, like any research, this study had some limitations. For instance, while ontologies offer powerful knowledge representation, their development and maintenance can be complex and resource intensive. Streamlining ontology creation processes and exploring automated techniques could mitigate this challenge. Moreover, ML models, though accurate (R^2 of 0.989 and RMSE of 1.200), might face challenges when applied to different bridge structures or under varying environmental conditions. It is essential to assess model generalization across diverse scenarios.

Moving forward, this research opens avenues for further exploration and implementation of ML and ontological techniques in the broader field of civil engineering and infrastructure management. The fusion of data-driven predictive models with semantic knowledge representation holds great promise for addressing complex structural challenges and optimizing decision-making processes across various domains. As technology continues to advance, we anticipate that these methods will play an increasingly integral role in ensuring the longevity and safety of critical infrastructure.

REFERENCES

[1] C. Wu, P. Wu, J. Wang, R. Jiang, M. Chen, and X. Wang, "Critical review of data-driven decision-making in bridge operation and maintenance," *Struct. Infrastruct. Eng.*, vol. 18, no. 1, pp. 47–70, Jan. 2022, doi: 10.1080/15732479.2020.1833946.

[2] X. Zhou and X. Zhang, "Thoughts on the development of bridge technology in China," *Engineering*, vol. 5, no. 6, pp. 1120–1130, Dec. 2019, doi: 10.1016/j.eng.2019.10.001.

[3] Y. Jiang, G. Yang, H. Li, and T. Zhang, "Knowledge driven approach for smart bridge maintenance using big data mining," *Autom. Construct.*, vol. 146, Feb. 2023, Art. no. 104673, doi: 10.1016/j.autcon.2022.104673.

[4] S. Hou, H. Li, and Y. Rezgui, "Ontology-based approach for structural design considering low embodied energy and carbon," *Energy Buildings*, vol. 102, pp. 75–90, Sep. 2015, doi: 10.1016/j.enbuild.2015.04.051.

[5] G. Ren, R. Ding, and H. Li, "Building an ontological knowledgebase for bridge maintenance," *Adv. Eng. Softw.*, vol. 130, pp. 24–40, Apr. 2019, doi: 10.1016/j.advengsoft.2019.02.001.

[6] C. Wu, P. Wu, J. Wang, R. Jiang, M. Chen, and X. Wang, "Ontological knowledge base for concrete bridge rehabilitation project management," *Autom. Construct.*, vol. 121, Jan. 2021, Art. no. 103428, doi: 10.1016/j.autcon.2020.103428.
[7] R. Li, T. Mo, J. Yang, S. Jiang, T. Li, and Y. Liu, "Ontologies-based domain knowledge modeling and heterogeneous sensor data integration for bridge health monitoring systems," *IEEE Trans. Ind. Informat.*, vol. 17, no. 1, pp. 321–332, Jan. 2021, doi: 10.1109/TII.2020.2967561.
[8] K. Liu and N. El-Gohary, "Ontology-based semi-supervised conditional random fields for automated information extraction from bridge inspection reports," *Autom. Construct.*, vol. 81, pp. 313–327, Sep. 2017, doi: 10.1016/j.autcon.2017.02.003.
[9] R. Assaad and I. H. El-Adaway, "Bridge infrastructure asset management system: Comparative computational machine learning approach for evaluating and predicting deck deterioration conditions," *J. Infrastruct. Syst.*, vol. 26, no. 3, 2020, Art. no. 04020032, doi: 10.1061/(asce)is.1943-555x.0000572.
[10] A. Sengupta, S. I. Guler, and P. Shokouhi, "Interpreting impact echo data to predict condition rating of concrete bridge decks: A machine-learning approach," *J. Bridge Eng.*, vol. 26, no. 8, 2021, Art. no. 04021044, doi: 10.1061/(ASCE)BE.1943-5592.0001744.
[11] K. Liu and N. El-Gohary, "Deep learning-based analytics of multisource heterogeneous bridge data for enhanced data-driven bridge deterioration prediction," *J. Comput. Civil Eng.*, vol. 36, no. 5, 2022, Art. no. 04022023, doi: 10.1061/(asce)cp.1943-5487.0001018.
[12] H. Jaafaru and B. Agbelie, "Bridge maintenance planning framework using machine learning, multi-attribute utility theory and evolutionary optimization models," *Autom. Construction*, vol. 141, Sep. 2022, Art. no. 104460, doi: 10.1016/j.autcon.2022.104460.
[13] Z. Liu, A. Sextos, A. Guo, and W. Zhao, "ANN-based rapid seismic fragility analysis for multi-span concrete bridges," *Structures*, vol. 41, pp. 804–817, Jul. 2022, doi: 10.1016/j.istruc.2022.05.063.
[14] Y. L. Jiang, G. Yang, and H. H. Song, "Dynamic optimization design of extradosed cable-stayed bridge under earthquake excitation," (in Chinese) *Eng. Mech.*, vol. 37, pp. 313–319, Jun. 2020, doi: 10.6052/j.issn.1000-4750.2019.04.S060.
[15] F. Saba and Y. Mohamed, "An ontology-driven framework for enhancing reusability of distributed simulation modeling of industrial construction processes," *Can. J. Civil Eng.*, vol. 40, no. 9, pp. 917–926, Sep. 2013, doi: 10.1139/cjce-2011-0489.
[16] J. Zhang, H. Li, Y. Zhao, and G. Ren, "An ontology-based approach supporting holistic structural design with the consideration of safety, environmental impact and cost," *Adv. Eng. Softw.*, vol. 115, pp. 26–39, Jan. 2018, doi: 10.1016/j.advengsoft.2017.08.010.
[17] X. Xu and H. Cai, "Semantic approach to compliance checking of underground utilities," *Autom. Construct.*, vol. 109, Jan. 2020, Art. no. 103006, doi: 10.1016/j.autcon.2019.103006.
[18] K. Liu and N. El-Gohary, "Bridge deterioration knowledge ontology for supporting bridge document analytics," *J. Construct. Eng. Manag.*, vol. 148, no. 6, 2022, Art. no. 04022030, doi: 10.1061/(ASCE)CO.1943-7862.0002210.
[19] J. Zhang and T. E. El-Diraby, "Social semantic approach to support communication in AEC," *J. Comput. Civil Eng.*, vol. 26, no. 1, pp. 90–104, 2012, doi: 10.1061/(ASCE)CP.1943-5487.0000130.
[20] M. Compton et al., "The SSN ontology of the W3C semantic sensor network incubator group," *J. Web Semantics*, vol. 17, pp. 25–32, Dec. 2012, doi: 10.1016/j.websem.2012.05.003.
[21] K. Banujan and S. Vasanthapriyan, "Bridge ontology architecture for knowledge management in bridge maintenance," in *Proc. Moratuwa Eng. Res. Conf. (MERCon)*, Jul. 2020, pp. 1–6, doi: 10.1109/MERCon50084.2020.9185335.
[22] S. Becker and N. Gebbeken, "An expert system for bridge inspection," in *Proc. Bridge Maintenance, Saf., Manag., Resilience Sustainability, 6th Int. Conf. Bridge Maintenance, Saf. Manag., Taylor Francis-Balkema*, 2012, pp. 1758–1763, doi: 10.1201/b12352-257.
[23] P. Zhang, C. Cui, C. Li, C. Zhang, and H. Liu, "Holistic design of energy pile bridge deicing system with ontology-based multiobjective decision making," *Frontiers Mater.*, vol. 8, Aug. 2021, Art. no. 710404, doi: 10.3389/fmats.2021.710404.
[24] Y. T. Chai and T.-K. Wang, "Evaluation and decision-making framework for concrete surface quality based on computer vision and ontology," *Eng., Construct. Archit. Manage.*, vol. 30, no. 10, pp. 4881–4913, Nov. 2023, doi: 10.1108/ECAM-01-2022-0064.
[25] J. France-Mensah and W. J. O'Brien, "A shared ontology for integrated highway planning," *Adv. Eng. Informat.*, vol. 41, Aug. 2019, Art. no. 100929, doi: 10.1016/j.aei.2019.100929.
[26] M. Fernández-López, A. Gómez-Pérez, and N. Juristo, "METHONTOLOGY: From ontological art towards ontological engineering," in *Proc. AAAI Conf. Artif. Intell.*, 1997. [Online]. Available: https://www.researchgate.net/publication/50236211
[27] M. Uschold and M. Gruninger, "Ontologies: Principles, methods and applications," *Knowl. Eng. Rev.*, vol. 11, no. 2, pp. 93–136, Jun. 1996, doi: 10.1017/s0269888900007797.
[28] R. Stammers. (2023). *Physical Infrastructure Research Contributing to Physical Infrastructure Research Contributing to Improved Resilience and Security Improved Resilience and Security*. [Online]. Available: https://stories.surrey.ac.uk/research-in-focus-physical-infrastructure/index.html
[29] H. Li and Y. Guo, "Estimating factors influencing the capacity of the wide-road-and-narrow-bridge section based on random forest," *Promet-Traffic Transp.*, vol. 35, no. 1, pp. 1–11, Feb. 2023, doi: 10.7307/ptt.v35i1.46.
[30] X. Liu, J. Qin, K. Zhao, C. A. Featherston, D. Kennedy, Y. Jing, and G. Yang, "Design optimization of laminated composite structures using artificial neural network and genetic algorithm," *Compos. Struct.*, vol. 305, Feb. 2023, Art. no. 116500, doi: 10.1016/j.compstruct.2022.116500.
[31] B. Kamiński, M. Jakubczyk, and P. Szufel, "A framework for sensitivity analysis of decision trees," *Central Eur. J. Oper. Res.*, vol. 26, no. 1, pp. 135–159, Mar. 2018, doi: 10.1007/s10100-017-0479-6.

YALI JIANG received the bachelor's and M.S. degrees from the Department of Civil Engineering, Dalian Maritime University, China, in 2016, where she is currently pursuing the Ph.D. degree. She is also a Visiting Student with Cardiff University. Her research interests include structure analysis and knowledge-based and artificial intelligence supported holistic and systematic decision making in the field of bridge maintenance.

HAIJIANG LI is currently the Chair of BIM for Smart Engineering. He is also the Head of the Research Group—Computational Mechanics and Engineering AI (CMAI), the Head of the Cross-Cutting Group—InSilico Digital Twins and Virtual Environments, and the Director of the Laboratory and M.Sc. Programme, BIM for Smart Engineering Centre, Engineering School, Cardiff University, U.K. For more than 20 years, he has been working toward a new generation smart computational engineering platform underpinned by super power computing, big data and machine learning, inter-linked domain knowledge, artificial intelligence algorithms and BIM-based standardization across sectors throughout project life cycle, to support large scale engineering data/information/knowledge processing, smart engineering optimization, and holistic decision making to help to achieve sustainable and resilient infrastructure system and built environment.

He is a fellow of the Institution of Civil Engineers (FICE) and British Computer Society (FBCS), the Vice Chair of EG-ICE—European Group of Intelligent Computing in Engineering, the Standards Committee Technical Executive—Building SMART, the Editor-in-Chief of *Construction Innovation*, and an Associate Editor of *Automation in Construction*.

GANG YANG received the bachelor's and M.S. degrees from Northeastern University, and the Ph.D. degree from the China University of Mining and Technology. Currently, he holds the position of a Professor and Ph.D. supervisor at the College of Transportation Engineering, Dalian Maritime University. Additionally, he serves as the Vice Chair of the Liaoning Provincial Society of Theoretical and Applied Mechanics. He had held a roles as the Head and Academic Leader of the Discipline of Road and Railway Engineering. His research interests include bridge engineering, structural engineering, tunnel engineering BIM technology research; virtual simulation design and calculation methods for bridge and tunnel structures; research on composite material reinforcement methods, structural stability calculations, and more.

CHEN ZHANG received the bachelor's degree from the School of Transportation Engineering, Huazhong University of Science and Technology, in 2004, and the master's degree from Dalian Maritime University, in 2007. He is currently a Senior Engineer and the Executive Director and the General Manager of ZJYY (Dalian) Bridge Underwater Inspection Company Ltd. He is also a Tutor of master's degree students with Dalian Maritime University. His research interests include the development of methods, equipment, tools for underwater bridge inspection, solving challenges related to underwater inspection and maintenance projects, and developing information systems for bridges.

KAI ZHAO received the Ph.D. degree in computational mechanics from Cardiff University.

He joined the group for structural health monitoring soon after graduation, where he applied his knowledge to develop machine learning surrogate models to predict the structural performance of bridge structures. Currently, he is a Postdoctoral Researcher with the University College Cork, focusing on road surface anomaly detection using machine leaning algorithms applied to vibration data recorded by smartphones. His research interests include numerical methods for structural analysis, design and optimization of composite structures, structural health monitoring of road and bridge structures, and applications of machine learning algorithms in engineering.