Four-Dimensional Fire Safety Knowledge–Enhanced Large Language Model for BIM Compliance Checking

Yicheng Zhao[1], Zhoupeng Wang[1], Xingbo Gong[2] and Xingyu Tao[1*]

[1] Dept. of Building and Real Estate, The Hong Kong Polytechnic University, Hong Kong 999077, China.
[2] Dept. of Civil and Environmental Engineering, The Hong Kong University of Science and Technology, Hong Kong 999077, China.
yicheng0220.zhao@connect.polyu.hk, zhou-peng.wang@connect.polyu.hk, xgongai@connect.ust.hk, xingyu.tao@polyu.edu.hk

Abstract

Fire safety compliance checking is essential for BIM-based design review, yet it remains largely manual in practice, leading to substantial time costs. Although automated checking approaches (e.g., NLP-based methods) have been explored, two major challenges remain: (1) in practice, operationalization is largely limited to semantic requirements, whereas clauses involving geometric and topological structures are difficult to execute; (2) automation remains limited because NLP-based methods often require manual mapping tables due to terminology mismatch and unclear condition parsing, and the extracted rules are not sufficiently stable. To address these issues, this paper proposes the Framework of Four-Dimensional Fire Safety Knowledge–Enhanced Large Language Model (4FKE-LLM). First, this paper constructs a four-dimensional fire safety knowledge graph—Semantic, Geometric, Topological, and Scenario, which organizes and structurally represents regulations across the four dimensions to support subsequent LLM-based compliance checking. Second, this paper proposes an LLM-based two-layer compliance checking mechanism. The first layer pre-filters regulatory and model information to identify applicable regulations and relevant objects, and the second layer performs dimension-specific regulation matching and reasoning across the four categories.

Experiments on a dormitory building case study show that the 4FKE-LLM framework achieves F1 = 0.863 and Accuracy = 0.885, and outperforms the two selected baselines on this case in terms of classification performance and checkable coverage.

Keywords –

BIM Compliance Checking; Fire Safety Regulations; Knowledge Graph; Four-Dimensional Rule Classification; Large Language Model

1 Introduction

Fire safety inspection is of paramount importance during the BIM design process. In the event of a fire, the evacuation of personnel, smoke control, fire compartmentalization, and firefighting efforts are directly affected by the spatial and component layout of the building [1]. If issues are not detected and corrected during the design phase, the subsequent costs and risks of rework during construction and operations can be significant, impacting both project timelines and safety [2, 3]. Additionally, BIM designs must comply with fire safety regulations [4]. When a project fails to meet key regulatory clauses, designers are required to modify their plans, further delaying the construction schedule. Therefore, conducting timely and reliable fire safety compliance checks during the BIM design phase is essential [1].

Currently, there are three main approaches for fire safety compliance checking in BIM. The first method is manual inspection. Designers or reviewers examine the drawings and models, cross-referencing with checklists, applying judgment based on experience, and performing spot calculations to verify compliance [2]. While this method can address complex cases and incorporate engineering expertise, it is time-consuming, costly, and subjective, with varying results depending on the individual examiner [3]. To overcome this limitation, researchers begin using rule engines. This approach decomposes regulatory clauses into explicit conditions and thresholds, then uses geometric calculations, database queries, or scripts to automatically assess the model [4, 5]. This approach is faster and more stable, but is highly dependent on the complete accuracy of the translation of the rules into a machine-readable format and the ability of the model to contain the required parameters and relationships [6, 7]. When clauses are complex or models lack essential parameters, the rules can fail to execute, reducing the scope of checks. The third method, based on natural language processing (NLP), attempts to automatically structure regulatory texts and align clauses with model fields for process-based checks [8, 9]. While this method reduces the need for manual clause breakdowns and organization, it still faces issues such as inconsistent terminology and unclear regulation decomposition [8]. These issues can lead to unstable clause-to-model matching, misclassification of requirements, and reduced checkable coverage in practical compliance workflows. More importantly, many fire safety clauses require spatial, component, and connectivity relationships, which are often insufficiently aligned with textual descriptions, limiting their reliability and verifiability [5-7].

In recent years, the capabilities of large language models (LLMs) have progressed rapidly. LLMs can better understand natural language clauses, normalize different expressions, and generate more coherent explanations [10]. This allows LLMs to address the shortcomings of traditional NLP in "semantic understanding" and "expressive rewriting". However, directly applying LLMs to compliance checking still leaves two key research gaps. First, LLMs can generate plausible but inaccurate conclusions and may produce inconsistent judgments across different reasoning rounds [3]. Second, fire safety compliance depends not only on semantics but also on geometric constraints, topological relationships, and attribute dependencies, such as "from where to where," "which doors connect to which corridors," "whether a compartment is independent," and "whether a component's fire resistance rating meets the requirement." If these relations are missing from the BIM model or are not structured into usable data, the LLM may understand the text but still cannot produce stable or explainable clause-level judgments. As a result, it tends to work better on semantic clauses while struggling with geometric, topological, and scenario logic [3]. Therefore, LLMs function here not as standalone inspectors, but as supporting reasoning modules. Accordingly, this study aims to enable stable, clause-level fire safety checking by integrating these models with structured four-dimensional rules and enriched BIM data.

Thus, this paper proposes the Framework of Four-Dimensional Fire Safety Knowledge–Enhanced Large Language Model (4FKE-LLM) to solve these issues. (Figure 1). First, 4FKE-LLM constructs a four-dimensional fire safety regulation knowledge graph (Semantic, Geometric, Topological, and Scenario), which organizes and structurally represents regulations across the four dimensions to support subsequent LLM-based compliance checking. Second, 4FKE-LLM implements an LLM-based two-layer compliance checking mechanism, the first layer pre-filters regulation and model information to identify applicable regulations and relevant objects, while the second layer performs dimension-specific regulation matching and reasoning across the four dimensions. Rather than simply replacing traditional NLP modules, the key contribution of 4FKE-LLM lies in integrating a four-dimensional rule representation with a two-layer LLM checking workflow to improve classification performance and checkable coverage.

2 Methodology

Framework of Four-Dimensional Fire Safety Knowledge–Enhanced Large Language Model (4FKE-LLM)

The 4FKE-LLM Framework integrates BIM and fire safety regulation rules to provide an efficient fire safety compliance checking method (Figure 1). The overall framework comprises three main components: (1) the development of the 4D Fire Safety Knowledge Graph (steps 1-3), (2) BIM fire safety information extraction and data enrichment (steps 4-6), and (3) LLM-based reasoning using "Four-Dimensional Rule Knowledge" and "Model Data" to generate a compliance checking report (steps 7-11).

The first component is the development of the 4D Fire Safety Knowledge Graph. This process involves denoising raw regulatory texts and categorizing them into the four dimensions of semantic, geometric, topological, and scenario-based rules (step 1), enabling the LLM to better understand the structure and content of regulations. The categorized rules are then converted into a knowledge graph (step 2) and stored in a data exchange format (step 3) that is easily understood by the LLM, facilitating subsequent queries.

The second component involves BIM fire safety information enrichment and extraction. Custom parameters are added to the BIM model to enrich the key data necessary for fire safety compliance checks (step 4), which are typically missing in the original BIM model. The relevant parameters are then extracted and converted into the Industry Foundation Classes (IFC) standard (step 5), ensuring interoperability between different platforms. The enriched IFC model is then translated into a data format suitable for the LLM (step 6). This ensures that the model can read and process BIM fire safety information and provides data support for checking.

The third component is LLM-based Two-layer fire safety compliance checking. User can submit requests through a pre-configured prompt at the query interface (step 7). After ingesting the 4D Fire-Safety Knowledge Graph and the BIM data (step 8), the framework performs two-layer checking. In the first layer, the LLM applies predefined filtering rules to identify applicable clauses and select relevant BIM objects (step 9). In the second layer, it conducts rule matching and reasoning over the

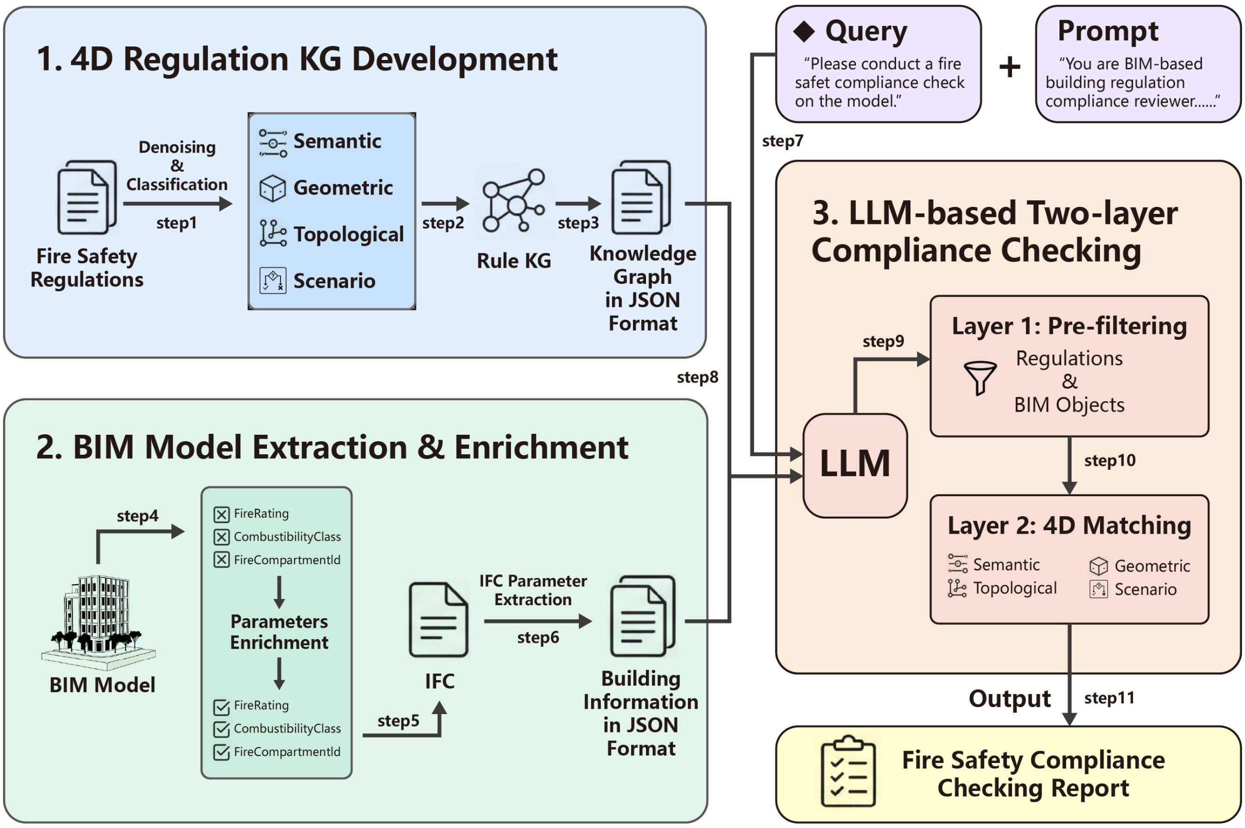


Figure 1: Framework of Four-Dimensional Fire Safety Knowledge–Enhanced Large Language Model

four-dimensional regulations using the combined knowledge graph and BIM model data to generate compliance reports (steps 10-11).

2.1 4D Rule Knowledge Graph Development

Figure 2 and Table 1 illustrate the classification of fire safety regulations into four types, each corresponding to different query designs. In this study, this four-type taxonomy is proposed as an operational representation for BIM-based fire safety compliance checking, and its broader applicability beyond the current fire regulations should be further tested in future work. The 4D Rule Knowledge Graph Development method aims to enhance the accuracy and automation of regulatory interpretation and compliance checking, reducing errors and omissions to achieve a more comprehensive and precise compliance evaluation.

Semantic Rules: Semantic regulations describe the relationships between regulations and the attributes that must be satisfied by specific components. For example, they specify the fire resistance rating required for fire doors.

$$(Regulation) - [HAS\_ATTRIBUTE\_REQ] -> (Component)$$

Geometric Rules: Geometric regulations focus on dimensional and distance constraints for components or spaces. These include minimum widths, maximum heights, and similar spatial metrics.

$$(Regulation) - [HAS\_GEOMETRIC\_REQ] -> (Component)$$

Topological Rules: Topological regulations define static relationships between components or spaces, such as connectivity, separation, or passage relations. These rules establish spatial connectivity between components using keywords such as "FROM" and "TO."

$$(Regulation) - [DEFINES] -> (TopoRule)$$

Scenario Rules: Scenario regulations define conditional logic in which a specific requirement becomes applicable only when a trigger condition is met. For instance, they specify that certain fire safety equipment must be installed when specific height requirements are met.

$$(Regulation) - [DEFINES\_LOGIC] \rightarrow (Scenario)$$

Semantic rules focus on components, attributes, and constraints, mapping clauses to component nodes and abstracting fire resistance ratings and other requirements into attributes and constraint nodes. Geometric rules primarily address dimensional constraints, while topological rules define static relations among spaces or components, like connectivity and isolation. In contrast, scenario rules focus on trigger–requirement logic, following a two-phase structure with "trigger conditions" and "requirements" to manage actions based on specific scenarios, enabling automatic filtering and reasoning.

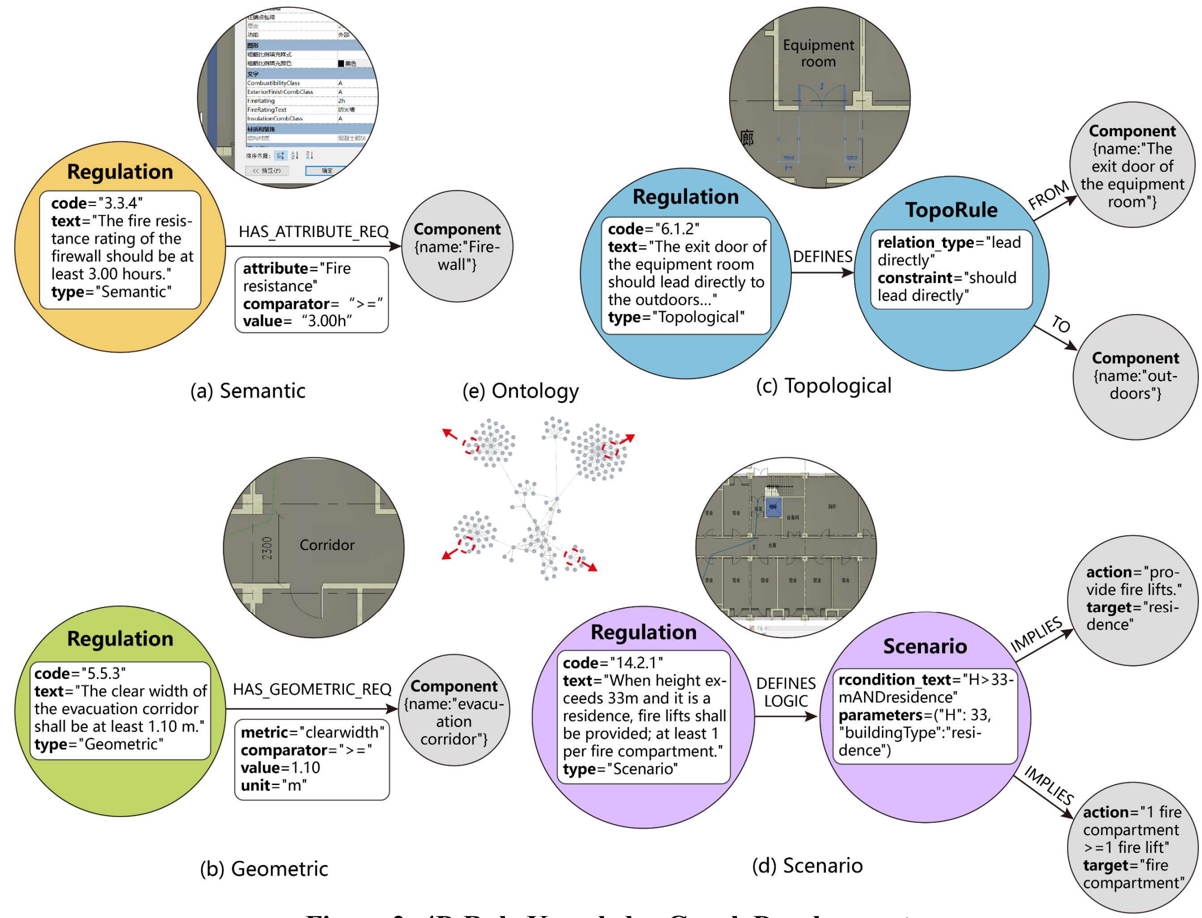


Figure 2: 4D Rule Knowledge Graph Development

Table 1

Node and Relationship Schema for the 4D Rule Knowledge Graph

Category	Node Design	Relationship Design
Semantic	Regulation: code, text, type = "Semantic". Component: name (the constrained component/object type).	**(Regulation)--[:HAS_ATTRIBUTE_REQ]→(Component)** • attribute: attribute name (e.g., fire resistance rating / combustibility) • value: required value (e.g., 2.00h / Class A / Grade A) • comparator: =, >=, <=, prohibition/permission, etc.
Geometric	Regulation: code, text, type = "Geometric". Component: name (component/space type).	**(Regulation)--[:HAS_GEOMETRIC_REQ]→(Component)** • metric: indicator (e.g., area / distance) • value: numeric value • unit: m, m², etc. • comparator: >=, <=, etc.
Topological	Regulation: code, text, type = "Topological". TopoRule: relation_type, constraint. Component: name (space/entity type).	**(Regulation)--[:DEFINES]→(TopoRule)** (the clause defines the rule) **(TopoRule)--[:FROM]→(Component_Source)** (source/endpoint) **(TopoRule)--[:TO]→(Component_Target)** (target/endpoint)
Scenario	Regulation: code, text, type = "Scenario". Scenario: condition_text, parameters. Requirement: action (required/prohibited action), target (affected object/scope).	**(Regulation)--[:DEFINES_LOGIC]→(Scenario)** (defines trigger conditions) **(Scenario)--[:IMPLIES]→(Requirement)** (requires compliance when the condition is met)

2.2 BIM Model Parameter Enrichment & Extraction

Figure 3 shows the BIM Model Parameter Enrichment and Extraction process. It adds key parameters for fire safety compliance checking, such as Fire Resistance Rating and Combustibility Class, to the BIM model so that most model elements contain information relevant to compliance checking. After parameter enrichment, the BIM model is exported in the IFC format. Fire safety attributes are bound to specific components using IFC standard property sets (Psets), enabling standardized cross-platform representation. Schedules and Mapping Tables are also exported to summarize parameters and standardize parameter mapping, naming, meaning, and units to ensure consistency across components and sources. These Mapping Tables are used for IFC/JSON field normalization and schema alignment, and they are not the manual regulation-to-model term mapping tables required by traditional rule-based compliance checking.

To ensure high-quality data input for compliance checks, this paper standardizes and denoises the data in two stages: "BIM to IFC" and "IFC to JSON." During the IFC export process, the Mapping Table is used to unify fire safety parameters' key names, meanings, and units, and essential relationships, such as those between components and rooms, are embedded directly into the data. During the conversion to JSON, the hierarchical and reference relationships in the IFC file are structured, and relevant attributes like GUIDs, component types, spatial contexts, and fire safety parameters are encapsulated in a directly accessible format, removing unrelated properties and complex geometric details. This process ensures that only the minimal necessary data required for the four types of rule checks remain in the final output.

The resulting model JSON is organized by "component—attribute—spatial context" and typically includes parameters such as element ID, IFC type, spatial context (e.g., storey, room), and fire safety-related attributes (e.g., Fire Rating). This JSON file serves as input for LLM-based fire safety compliance checks. It enables direct retrieval and reference, which improves the stability and efficiency of compliance checking.

2.3 LLM-based Two-layer Compliance Checking

The third component is LLM-based Two-layer fire safety compliance checking. Users submit requests through a pre-configured prompt via the query interface, specifying the checking scope (e.g., building type, floor range, or target spaces). After the 4D Fire Safety Knowledge Graph and enriched BIM model data are provided as inputs, the framework performs two-layer checking. In Layer 1, the LLM filters applicable clauses and retrieves the relevant BIM objects. In Layer 2, it evaluates each filtered clause against the corresponding BIM evidence under the semantic, geometric, topological, or scenario category and then generates the compliance report.

3 Validation of the Framework

Figure 4 illustrates the validation process for the framework, which was conducted on a dormitory building case study. The regulations selected for testing include the "General Fire Safety Code for Buildings" GB 55037-2022 [11], the "Fire Safety Code for Interior Building Decorations" GB 50222-2017 [12], and the "Building Fire Safety Code" GB 50016-2014 (2018 edition) [13]. First, the regulatory text was denoised, and

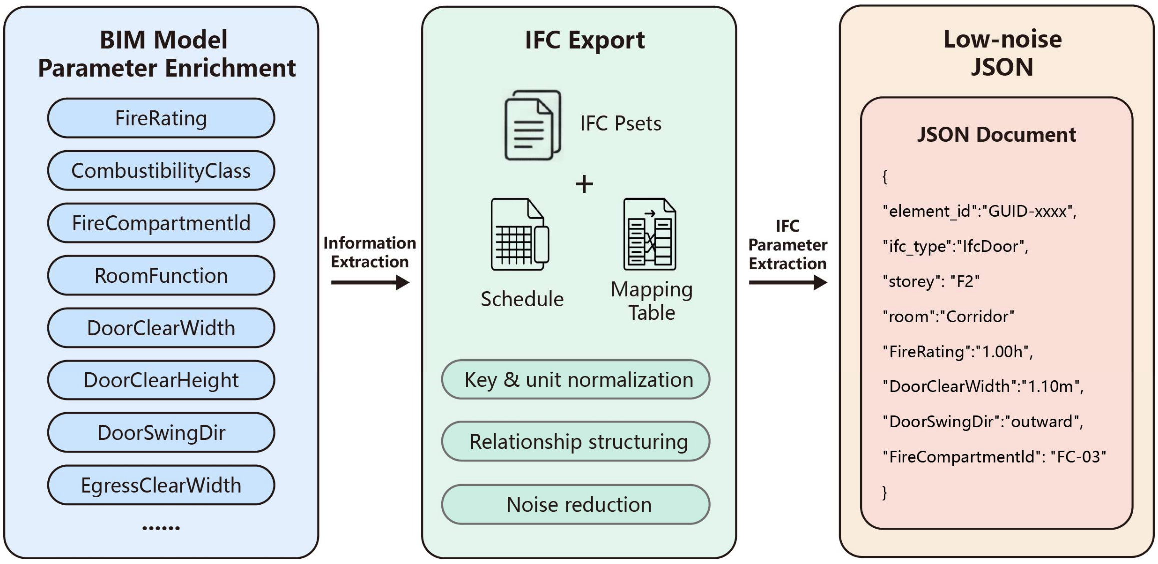
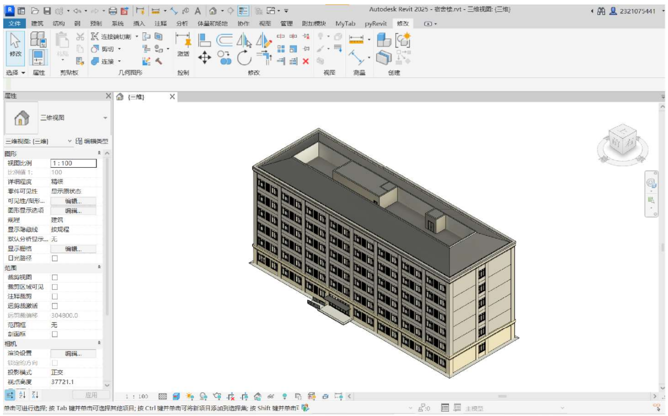


Figure 3: BIM Model Parameter Enrichment & Extraction



(a)Revit Model

Role:

BIM-based building-code compliance reviewer (semantic/attribute checks).

Goal:

For each clause, output COMPLIANT / NON-COMPLIANT using provided evidence.

Scope:

Dormitory; Levels F1–F2 only.

Ignore MEP/operation/signage/soft decoration/site checks …

Inputs:

(A) Semantic Regulation JSON (B) BIM JSON

Rules:

- Filter in applicable clauses; exclude irrelevant building types …
- Clause type: conditional → verify trigger then check

Output:

{clause type, code, original text, decision...}

Reply "Ready", then I will send A and B.

(b)Checking Prompt

Fire Safety Compliance Checking Report

Category	Code	Original Text	4FKE LLM Framework	Result	LLM-RAG Baseline	Result	LLM Only Baseline	Result
Semantic	6.4.4-2	For buildings with a height greater than 100 m, the inspection door on the shaft well wall shall be a Class A fire door.	Compliant	√	Compliant	√	—	×
Geometric	6.4.5-1	The height of the handrail shall not be less than 1.10 m, and the clear width of the stair flight shall not be less than 0.90 m.	Non-compliant	√	Compliant	×	—	×
Scenario	6.4.7	Evacuation stairs shall not use spiral stairs or fan-shaped treads, unless specific occupant-number and size conditions are met.	Non-compliant	√	Compliant	×	Compliant	×
Geometric	6.4.8	The clear width between stair flights of public evacuation stairs shall not be less than 150 mm.	Non-compliant	√	Non-compliant	√	Non-compliant	√
Topological	6.5.1-5	Where fire doors are provided near building movement joints, they shall be arranged on the side with more stairs, and it shall be ensured that when the fire door is opened, the door leaf does not cross the movement joint.	Compliant	√	Non-compliant	×	—	×
Semantic	6.5.3-3	The combustibility class of the interior finish materials for the ceiling, walls, and floor in the fire elevator anteroom or shared anteroom shall be Class A.	Compliant	√	Compliant	√	Compliant	√
Semantic	6.5.6-3	When passenger transfer lobbies and transfer passages are located underground or semi-underground, interior finish materials in the room shall not use combustible materials, asbestos products, glass fiber products, or plastic	Compliant	√	Compliant	√	—	×
Geometric	7.1.11	The handrail height of outdoor evacuation stairs shall not be less than 1.10 m, and the inclination angle shall not be greater than 45°.	Non-compliant	√	Compliant	×	—	×
Scenario	7.1.14	Industrial and civil buildings with a height greater than 100 m shall be provided with refuge floors.	Compliant	√	Compliant	√	—	×

......

(c)Checking Report

Figure 4: Validation of the Framework

tables were transcribed into natural language, yielding a total of 554 regulatory clauses. The rules were then classified into semantic, geometric, topological, and scenario categories and converted into Cypher queries, which were imported into Neo4j to construct the regulatory knowledge graph [14]. Structured rule data was exported as JSON for subsequent fire safety compliance checks using the LLM.

A BIM model of the dormitory building was created using Revit 2025 (Figure 4(a)), exported in IFC4 format to ensure that spatial boundary relationships and key fields were extractable. The model was then converted into structured JSON data, containing entity types, spatial hierarchies, connectivity relationships, and trigger conditions. During the compliance checking phase, the ChatGPT 5.2 Thinking model was used [15], with a pre-configured prompt (Figure 4(b)) that submitted the "4D Fire Safety Rule JSON + Model JSON" as input. This model was selected because the task requires long-context input handling and multi-step clause-level reasoning over rule JSON and model JSON. As the model is proprietary and may evolve over time, the reported results should be interpreted as a system-level evaluation under a fixed prompt setting rather than a strictly version-stable benchmark. First, the applicable clauses for the dormitory building were filtered, resulting in 122 relevant regulations. These clauses were then validated according to the four types of rules, and a compliance report was generated in Compliant/Non-compliant form (Figure 4(c)).

There are two baselines used for comparison in experiments. First, the LLM-RAG baseline takes the raw regulatory text together with the structured BIM model JSON as input. Second, the LLM-only baseline provides only the structured model JSON and asks the LLM to determine compliance without explicit rule inputs. The generated report is shown in Figure 4(c). For evaluation, each applicable clause-level decision is treated as a binary classification task, where the manually verified result is the ground truth label (Actual) and the model output is the predicted label (Pred). We report the confusion matrix and compute Accuracy, Precision, Recall, and F1, where Precision measures the reliability of non-compliance predictions, Recall measures the coverage of true non-compliant clauses, and F1 summarizes the trade-off between them. The metrics are defined in Eqs. (1) – (4).

$$accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$precision = \frac{TP}{TP + FP} \quad (2)$$

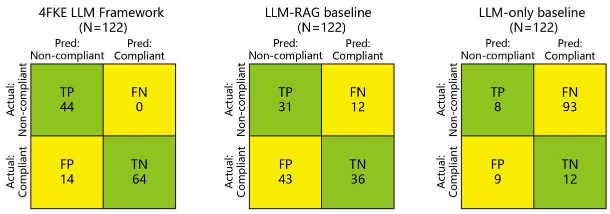
$$recall = \frac{TP}{TP + FN} \quad (3)$$

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (4)$$

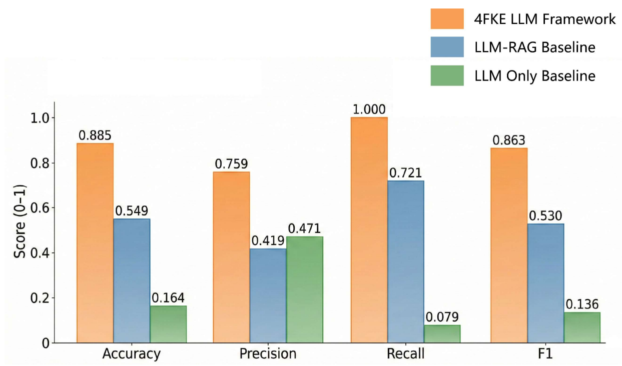
These four metrics are appropriate provided that the positive class is clearly defined. In this study, “Non-compliant” is treated as the positive class, so TP/FP/FN are counted with respect to non-compliance. Under this definition, Precision and Recall directly reflect the two key risks in fire-safety checking: false alarms (FP) and missed violations (FN). We choose these metrics because they capture complementary aspects of clause-level compliance checking: (1) Accuracy reports overall correctness, (2) Precision reflects the cost of false positives (unnecessary manual review or rework), (3) Recall reflects the safety-critical cost of false negatives (missed violations), and (4) F1 provides a balanced summary, avoiding misleading conclusions when class distributions are imbalanced.

Figure 5 summarizes the results (N = 122). In this context, N denotes the total number of clause-level compliance decisions evaluated, that is, the number of applicable clauses assessed across the case. The proposed Framework of Four-Dimensional Fire Safety Knowledge–Enhanced Large Language Model (4FKE-LLM) shows the best performance among the three evaluated settings, with TP = 44, TN = 64, FP = 14, FN = 0, yielding Accuracy = 0.885, Precision = 0.759, Recall = 1.000, and F1 = 0.863. This indicates zero missed violations (FN = 0) while maintaining a low false-positive rate. In contrast, the LLM-RAG baseline shows both higher false positives and false negatives (TP = 31, TN = 36, FP = 43, FN = 12), with reduced scores (Accuracy = 0.549, Precision = 0.419, Recall = 0.721, F1 = 0.530). The LLM-only baseline performs the worst, exhibiting severe missed violations (TP = 8, TN = 12, FP = 9, FN = 93; Accuracy = 0.164, Precision = 0.471, Recall = 0.079, F1 = 0.136), suggesting that providing model data alone without explicit regulatory rules does not yield stable clause-level compliance determinations.

In terms of practical expectations for safety-critical fire-safety compliance checking, this study adopts a recall-first evaluation criterion. It aims for Recall as close to 1 as possible (ideally FN = 0, or at least Recall ≥ 0.95) while keeping false positives within a manageable range (e.g., Precision ≥ 0.70 or a low FP rate), and it uses an overall target such as F1 ≥ 0.80 as a pragmatic usability threshold. This recall-first stance is consistent with prior automated fire code–compliance checking studies, which report very high precision/recall performance and, in some settings, argue that Recall is more critical than Precision [9, 16]. Under these criteria, 4FKE-LLM already falls within the expected range (Recall = 1.000, Precision ≈ 0.76, F1 ≈ 0.86), aligning with the risk preference of fire-safety review.



(a) Confusion Matrices



(b)Metrics Comparison

Figure 5: Comparison Result

4 Conclusion

This paper addresses the practical needs of BIM-based fire safety compliance checking by integrating a four-dimensional rule representation, BIM data enrichment, and a two-layer LLM checking workflow for clause-level reasoning. The Framework of Four-Dimensional Fire Safety Knowledge–Enhanced Large Language Model (4FKE-LLM) constructs regulatory knowledge graphs using a four-dimensional classification method and exports callable structured data. BIM model parameters are enriched and converted into structured data that includes attributes, spatial structures, and key relationships. In the dormitory building case study, the framework outperformed the two selected baselines, especially in recall and overall F1. These results suggest that the framework can improve clause-level fire safety compliance checking for the evaluated case. However, the current validation is limited to one dormitory building and 122 applicable clauses, so broader testing across more building types and regulatory contexts is still needed to examine external validity. Furthermore, certain clauses still rely on finer-grained geometric calculations, and incomplete IFC exports or parameter enrichment may lead to situations that cannot be checked. Future work will focus on expanding the rule coverage and the clause template library. It will also improve IFC-based attribute mapping and the geometric/topological evidence computation modules to enhance cross-project compatibility and robustness.

Acknowledgement

This research is supported by the Hong Kong PolyU (UGC) Start-up Fund for New Recruits (P0056204) and Zhejiang Provincial "Pioneer" and "Leading Goose" Science and Technology Plan Project (2026C04040)

Reference

[1] S. Fischer, C. Schranz, H. Urban, and D. Pfeiffer, "Automation of escape route analysis for BIM-based building code checking," *Automation in construction,* vol. 156, p. 105092, 2023, doi: 10.1016/j.autcon.2023.105092.

[2] R. Amor and J. Dimyadi, "The promise of automated compliance checking," *Developments in the built environment,* vol. 5, p. 100039, 2021, doi: 10.1016/j.dibe.2020.100039.

[3] O. Iversen and L. Huang, "Leveraging large language models for BIM-based automated compliance checking," *Automation in construction,* vol. 182, p. 106707, 2026, doi: 10.1016/j.autcon.2025.106707.

[4] A. S. Ismail, K. N. Ali, N. A. Iahad, M. A. Kassem, and N. T. Al-Ashwal, "BIM-Based Automated Code Compliance Checking System in Malaysian Fire Safety Regulations: A User-Friendly Approach," *Buildings (Basel),* vol. 13, no. 6, p. 1404, 2023, doi: 10.3390/buildings13061404.

[5] J. Peng and X. Liu, "Automated code compliance checking research based on BIM and knowledge graph," *Scientific reports,* vol. 13, no. 1, pp. 7065-12, 2023, doi: 10.1038/s41598-023-34342-1.

[6] A. Li, P. K.-Y. Wong, X. Tao, J. Ma, and J. C. P. Cheng, "An interactive system for 3D spatial relationship query by integrating tree-based element indexing and LLM-based agent," *Advanced engineering informatics,* vol. 66, p. 103375, 2025, doi: 10.1016/j.aei.2025.103375.

[7] X. Tao *et al.*, "Smarter smart contracts for automatic BIM metadata compliance checking in blockchain-enabled common data environment," *Advanced engineering informatics,* vol. 62, p. 102627, 2024, doi: 10.1016/j.aei.2024.102627.

[8] J. Zhang and N. M. El-Gohary, "Semantic NLP-Based Information Extraction from Construction Regulatory Documents for Automated Compliance Checking," *Journal of computing in civil engineering,* vol. 30, no. 2, 2016, doi: 10.1061/(ASCE)CP.1943-5487.0000346.

[9] Y. Wang, Y. Liu, H. Cai, J. Wang, and X. Zhou, "An Automated Fire Code Compliance Checking Jointly Using Building Information Models and Natural Language Processing," *Fire (Basel, Switzerland),* vol. 6, no. 9, p. 358, 2023, doi: 10.3390/fire6090358.

[10] S. Madireddy *et al.*, "Large Language Model-Driven Code Compliance Checking in Building Information Modeling," *Electronics (Basel),* vol. 14, no. 11, p. 2146, 2025, doi: 10.3390/electronics14112146.

[11] *General Fire Safety Code for Buildings (GB 55037-2022)*, Ministry of Housing and Urban-Rural Development of the PRC, Beijing, 2022.

[12] *Code for Fire Prevention in Design of Interior Decoration of Buildings (GB 50222-2017)*, Ministry of Housing and Urban-Rural Development of the PRC, Beijing, 2017.

[13] *Code for Fire Protection Design of Buildings (GB 50016-2014, 2018 Edition)*, Ministry of Housing and Urban-Rural Development of the PRC, Beijing, 2018.

[14] I. Neo4j. "Neo4j Graph Database (Web Version)." https://console.neo4j.io/ (accessed January 20, 2026).

[15] OpenAI. "ChatGPT (Version 5.2 Thinking model)." https://chat.openai.com (accessed January 20, 2026).

[16] J. Wu, X. Xue, and J. Zhang, "INVARIANT SIGNATURE, LOGIC REASONING, AND SEMANTIC NATURAL LANGUAGE PROCESSING (NLP)-BASED AUTOMATED BUILDING CODE COMPLIANCE CHECKING (I-SNACC) FRAMEWORK," *Journal of information technology in construction,* vol. 28, pp. 1-18, 2023, doi: 10.36680/j.itcon.2023.001.