

LLM-Empowered Sensor Data Analytics and Interactive Digital Twin for Building Facility Management

Tianlun Yang^{1*}, Rong Chen², Haojie Tan³, Georgios Kapogiannis⁴, Byung-Gyoo Kang⁵

¹Department of Intelligent Interaction Design, NingboTech University, Ningbo 315040, Zhejiang, China

²Ningbo International Investment Consulting Co., Ltd., Ningbo 315040, Zhejiang, China

³Zhejiang Wanli Education Group, Ningbo 315040, Zhejiang, China

⁴Warwick Manufacturing Group, University of Warwick, Coventry CV4 7AL, West Midlands, UK

⁵Department of Civil Engineering, University of Nottingham Ningbo China, Ningbo 315040, Zhejiang, China

*Corresponding author: Tianlun Yang, tianlun.yang@nbt.edu.cn

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: This study proposes an LLM-empowered edge-cloud digital twin architecture for intelligent FM. Edge servers acquire and pre-process sensor data, with Flask-based cloud middleware aggregating into hybrid data storage. The LLM performs semantic reasoning to generate operational decisions, relaying insights in real time to a Unity-engineered 3D visualization environment. BIM semantic parsing segregates geometric from non-graphical attributes, integrating multi-source data at room-floor granularity to establish a “perception-analysis-decision-execution” closed loop, thereby enhancing situational awareness and decision support for facility managers.

Keywords: Digital twin; LLM; Facility management

Online publication: March 31, 2026

1. Introduction

Facility management (FM) faces critical digital transformation challenges. While Internet of Things (IoT) technologies enable massive sensor networks for monitoring equipment states and environmental parameters, facility managers struggle with fragmented, heterogeneous data streams that hinder real-time situational awareness, resulting in delayed maintenance decisions and escalating operational costs. Conventional FM systems lack intelligent multi-source data parsing capabilities and natural human-machine interfaces, failing to translate raw sensor data into actionable knowledge. This study builds upon the authors’ established digital twin and facility management frameworks, yet these solutions still have room for improvement in real-time semantic parsing and dynamic decision support^[1-3]. While large language models (LLMs) offer powerful semantic reasoning capabilities to overcome the “data-rich yet information-poor” predicament and digital twin (DT) technology enables

immersive state awareness, existing studies predominantly treat them as separate technology stacks. This research introduces an LLM-empowered edge-cloud collaborative architecture that deeply integrates both technologies, enabling intelligent sensor data parsing and immersive visual interaction to deliver dynamic, interpretable decision support for enhanced building operational intelligence.

2. Literature review

BIM has evolved into a comprehensive infrastructure for FM lifecycle operations, with empirical studies validating its role in mitigating information fragmentation during construction-to-operations handover and enhancing predictive maintenance through structured asset repositories ^[4,5]. Subsequent research has focused on semantic enrichment, including ontology-mediated approaches to harmonize BIM and FM knowledge domains and spatial indexing methodologies for room-scale asset management, collectively underscoring BIM's dependence on data interoperability standards such as IFC ^[6,7]. Recent scholarship has further investigated BIM-IoT convergence for real-time FM analytics, demonstrating that fusing static BIM geometry with dynamic sensor streams enhances space utilization diagnostics ^[8]. However, persistent challenges in automated semantic alignment between heterogeneous schemas and operational data formats continue to constrain scalability, necessitating intelligent mediation mechanisms to automate data reconciliation across the built environment lifecycle.

The convergence of DT with BIM has shifted static documentation toward dynamic lifecycle management, grounded in Grieves and Vickers' (2017) theoretical foundation of synchronized virtual-physical counterparts and further articulated through BIM's evolution from "as-designed" to "as-operated" models via continuous IoT updates ^[9,10]. Technical mechanisms for synchronization, including cloud-based middleware for bidirectional data exchange and Unity-based immersive interfaces for interactive decision-support, have attracted substantial attention ^[11,12]. However, data governance challenges and computational intensity constrain scalable implementation, as schema inconsistencies between BIM, IoT, and FM workflows present critical barriers to automated synchronization, and real-time geometry updates impose prohibitive cognitive burdens for large portfolios ^[13]. These limitations motivate integrating LLMs to automate semantic mediation and cognitive reasoning within DT-BIM ecosystems, thereby reducing manual intervention and enhancing lifecycle decision-support capabilities.

The integration of LLMs into DT ecosystems marks a paradigm shift from predictive analytics to cognitive decision support, with research demonstrating LLMs' capacity to mediate semantic heterogeneity by translating unstructured operational logs into structured maintenance insights, serve as interpretability layers converting sensor anomalies into diagnostic narratives, and deploy retrieval-augmented generation to ground outputs in real-time sensor streams ^[14-16]. Within architecture, engineering, and construction, emergent applications include natural language interfaces for HVAC performance querying and automated regulatory compliance verification ^[17,18]. However, latency constraints, model opacity in safety-critical contexts, and domain adaptation overhead constitute persistent barriers, necessitating purpose-built LLM-DT architectures tailored for facility management intelligence.

3. Research design

The research question is how operational inefficiencies in building asset management can be mitigated through enhanced real-time situational awareness. The aim of this research is to develop an LLM-driven DT system that integrates real-time sensor analytics with immersive 3D visualization to enhance decision-making support. The objectives of the research are as follows:

- (1) Information bottleneck analysis of building asset DT systems;
- (2) LLM-based semantic parsing and reasoning methods for sensor data;
- (3) Conceptual design an LLM-DT coupled interactive visualization architecture.

This study integrates literature review with conceptual modeling. For objective (1), a critical analysis of technical bottlenecks and architectural limitations in existing building DT systems for asset operation and maintenance is conducted. For objective (2), LLM application patterns in time-series data analysis and knowledge reasoning are synthesized to delineate technical pathways tailored to building operation and maintenance scenarios. For objective (3), a three-tier edge-cloud-terminal architecture is conceptualized and assessed through simulation-driven proof-of-concept against typical FM tasks, establishing theoretical foundations for subsequent engineering implementation.

4. Conceptual design intervention

This study proposes an LLM-empowered DT Framework for building facility management, achieving digital mirroring, semantic data management, cross-layer visualization, and intelligent closed-loop control through an edge-cloud collaborative system integrating environmental monitoring, occupancy recognition, and equipment tracking. The sensing infrastructure comprises the following four modalities:

- (1) Envelope status: Contact sensors at doors and windows capturing binary states as environmental boundary conditions;
- (2) Ambient conditions: Temperature, humidity, and illuminance sensors providing baselines for HVAC and lighting control;
- (3) Occupancy detection: Computer vision determining space utilization via edge or cloud inference for demand-driven energy optimization;
- (4) Equipment feedback: Electrical sensors on air handling units and luminaires monitoring operational states to close the control loop.

As illustrated in **Figure 1**, the system comprises three core modules: intelligence, visualization, and interaction. Multi-source sensor data are pre-processed at edge nodes and uploaded to the cloud, where LLMs perform semantic reasoning and generate decisions fed back to the Unity-based DR in real time. Users interact via cross-platform terminals to obtain decision support and issue control commands. The system adopts a BIM semantic parsing strategy that decouples graphical from non-graphical attributes, integrating LLM outputs through room-floor spatial indexing to enable spatially-oriented querying and operational decision-making for FM personnel.

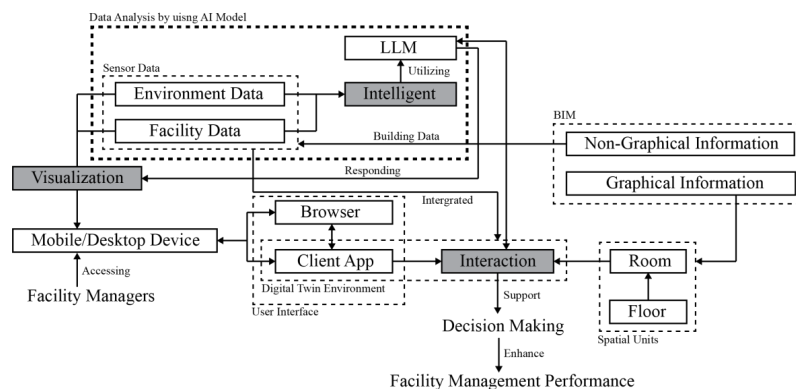


Figure 1. LLM-driven digital twin architecture for human-centric interactive decision support.

As illustrated in **Figure 2**, the system employs a hierarchical distributed architecture enabling real-time synchronization between physical and digital spaces. The edge layer performs room-level data aggregation and low-latency preprocessing via multi-type sensors and embedded platforms. The cloud layer ingests edge data through microservices, utilizing distributed storage and LLM-driven diagnostic insights. The application layer renders real-time 3D visualization through game-engine clients, allowing FM personnel to execute queries, anomaly localization, and remote control, thereby closing the “sensing-analysis-decision-execution” loop.

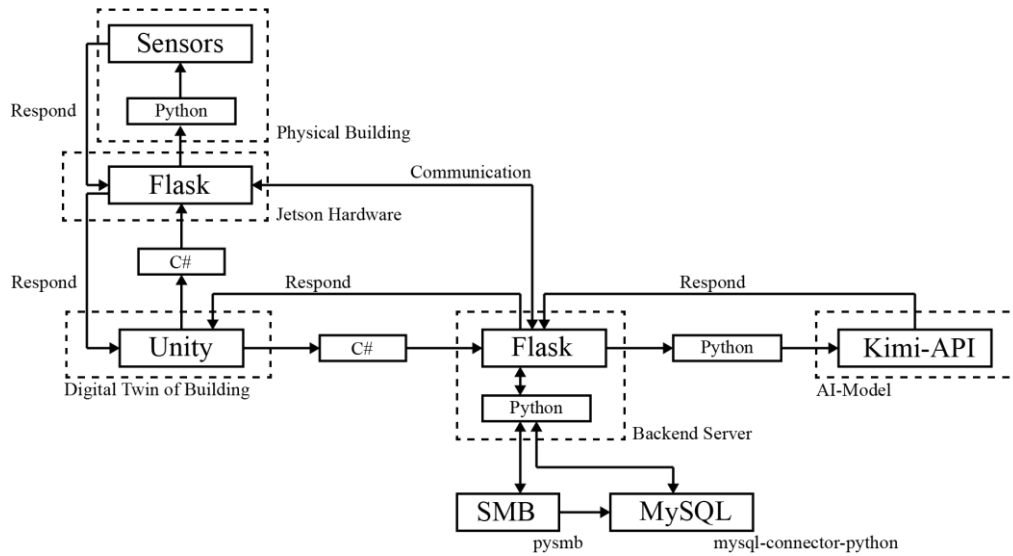


Figure 2. Technical roadmap of edge-cloud orchestrated DT for real-time data-driven interaction.

As illustrated in **Figure 3**, the hardware deployment adopts a three-tier architecture optimized for data scale, computational load, and latency. The field device layer comprises sensor networks and edge servers for data collection and lightweight pre-processing, enabling low-latency local response. The cloud server layer hosts high-performance computing for application services, databases, and LLM inference. The user terminal layer stores large-volume static assets (e.g., digital twin models) locally while dynamically fetching lightweight data via cloud APIs, implementing a hybrid architecture of local rendering and cloud linkage.

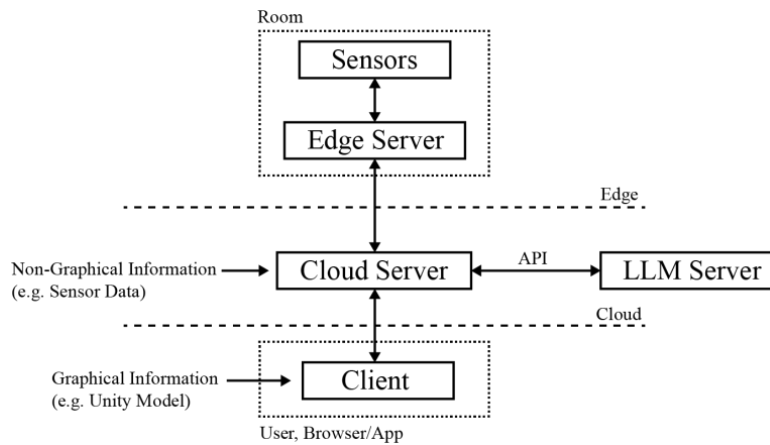


Figure 3. Tiered hardware infrastructure for edge-cloud digital twin systems.

5. Scenario-based validation via simulation

This study employs a “simulation-driven proof-of-concept” methodology to evaluate technical feasibility under reduced hardware dependency, examining edge-cloud data stability, multi-source semantic alignment, and command synchronization performance to optimize architectural design and mitigate engineering risks.

The occupant-unaware energy waste phenomenon during commercial building lunch breaks (12:00 to 13:30), where air-conditioning and lighting remain active with open doors and windows after occupant departure. This scenario offers as follows:

- (1) Universality: A prevalent operational pain point ensuring generalizability;
- (2) Moderate scale: Multi-point coordination across 20 offices testing scalability without excessive complexity;
- (3) Data integrity: Encompassing occupancy, envelope status, and equipment feedback for comprehensive edge-cloud and intelligent decision validation;
- (4) Quantifiable outcomes: Air-conditioning-lighting-door/window linkages readily formalized as ECA (event-condition-action) rules for energy-saving and latency assessment.

For deductive architecture, the standard floor deploys heterogeneous sensors (door/window status, environmental parameters, occupancy detection, equipment feedback) with room-level edge aggregation and preprocessing (**Figure 4**). The cloud layer implements an ECA engine, where occupancy absence combined with open envelope status triggers equipment shutdown, while environmental deviations invoke adaptive regulation strategies.

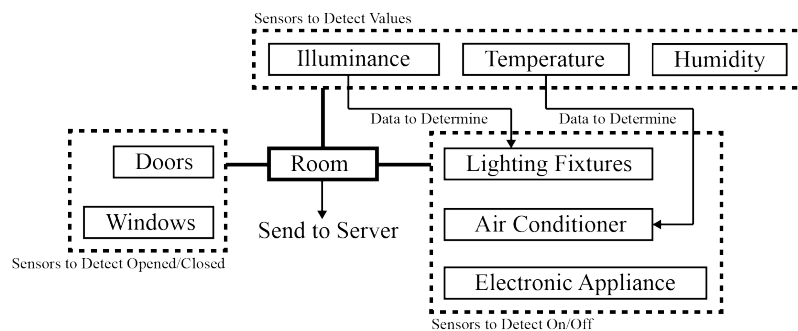


Figure 4. Edge-cloud orchestrated multi-source sensing and control architecture at room scale

For the deductive process, the observation window spans 12:00 to 13:30. Assuming 10 offices satisfy the “unoccupied and window-open” condition at 13:00, the system completes control command generation, digital twin state synchronization, and physical equipment actuation within a short timeframe, achieving improved response efficiency compared to conventional manual inspection. At 13:30, the deduction concludes; time-series data are archived in the database, supporting the LLM analysis module in identifying high-frequency energy waste spaces and generating targeted inspection recommendations, thereby realizing the paradigm shift from passive response to proactive prevention.

The system adopts a hierarchical decision architecture. The rule layer employs occupancy and door/window status as inputs, generating high-priority equipment on/off commands through deterministic logic to ensure real-time performance; the condition layer generates adaptive optimization strategies for equipment operating parameters based on dynamic comparison between environmental parameters and comfort thresholds. This architecture balances the deterministic latency of rule engines with the flexible regulation of multi-parameter coordination, forming a closed-loop control mechanism of “environment-occupant-equipment” (**Figure 5**).

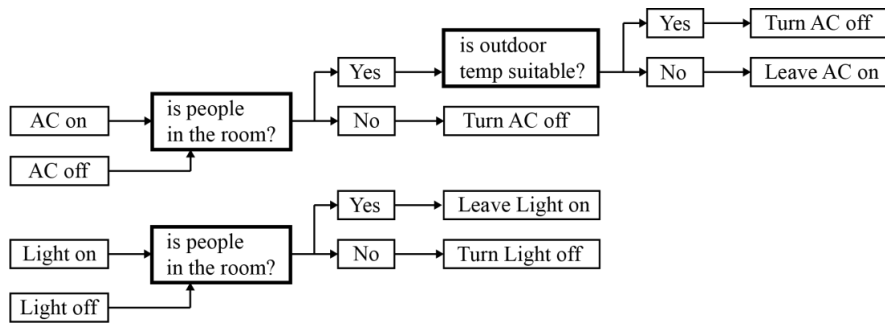


Figure 5. Hierarchical control decision framework for sensor-equipment linkage.

6. Discussion

The core contribution is an LLM-Digital Twin architecture enabling real-time semanticization and spatial visualization through edge-cloud computing and BIM parsing, characterized by three integrations as outlined:

- (1) Data integration: Unified semantic encoding of multi-source sensor data at the edge;
- (2) Reasoning integration: Complementing deterministic rules with open-ended LLM inference;
- (3) Interaction integration: Reducing cognitive load via 3D digital twins and natural language interfaces.

This advances existing research by employing LLM as an active reasoning engine, shifting from “human adapting to system” to “system adapting to human” while addressing building operations’ high dynamics and spatial embeddedness absent in structured industrial environments (**Figure 6**). The edge-cloud architecture and lightweight BIM parsing provide a reusable pathway for resource-constrained, real-time intelligent facility management.

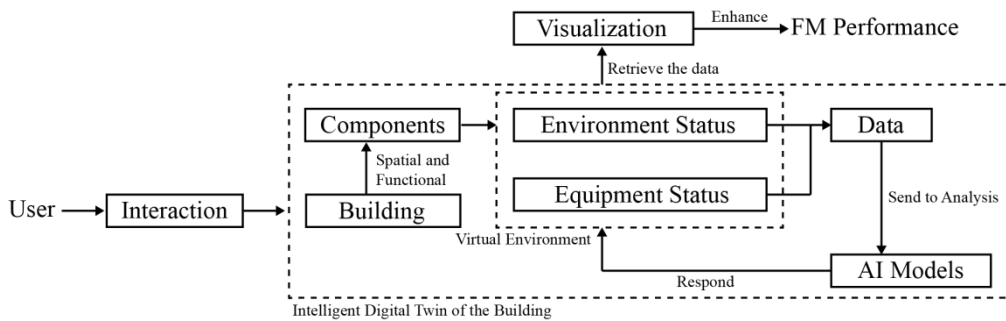


Figure 6. Impact of DT-LLM integration on facility management efficiency and decision quality.

Theoretically, this study addresses the cognitive gap in human-machine collaborative decision-making. Traditional FM systems demand cross-domain expertise (HVAC, electrical, data analytics) from operators, whereas this architecture transforms complex raw data into actionable knowledge through LLM-enabled semantic abstraction and visualization mapping, enabling FM personnel to focus on high-order judgment rather than low-level data interpretation. This cognitive outsourcing mechanism extends the decision-making capability boundaries of non-expert users, providing empirical reference for human-machine symbiosis in intelligent infrastructure management.

Practically, the edge-cloud hierarchical design balances computational efficiency with deployment costs, where edge nodes (embedded platforms) handle latency-sensitive preprocessing and control command issuance, while the cloud layer focuses on non-real-time LLM inference and historical data analytics. This division of labor

aligns with the economic constraints of existing building intelligence retrofits. Conceptual deduction demonstrates that the system can achieve a quick response to the management of the office, improving efficiency compared to conventional manual inspection, indicating potential for scalable deployment.

7. Conclusion

This study addresses the core contradiction between data visualization and cognitive decision-making in building facility management by constructing an LLM-empowered edge-cloud collaborative digital twin architecture, and evaluates its technical feasibility through simulation-driven proof-of-concept. Key findings include as listed:

- (1) Edge-cloud hierarchical computing effectively balances heterogeneous demands for real-time control and intelligent reasoning, achieving second-level responsive closed-loop equipment management;
- (2) LLM semantic reasoning transforms multi-source sensor data into interpretable operation and maintenance recommendations, significantly reducing the professional cognitive load of facility management personnel;
- (3) The integration of BIM semantic indexing and spatial visualization supports cross-scale information retrieval from building to equipment, enhancing spatiotemporal contextual awareness for operation and maintenance decision-making.

8. Limitation and future research

This study remains at the proof-of-concept stage, with three primary limitations constraining its current validity:

- (1) Limited scenario coverage, as deductions derive solely from a single office building typology, leaving the generalizability of large language models (LLMs) in complex operational environments such as industrial or healthcare facilities untested;
- (2) Untested technical robustness regarding edge-cloud communication stability and sensor fault tolerance under real-world network conditions;
- (3) Insufficient empirical validation, where projected energy savings and operational efficiencies rely exclusively on simulation data without comparative baselines against manual operation modes.

Future research will address these gaps through longitudinal field deployment across diverse building typologies, domain-specific LLM fine-tuning using operational feedback, and controlled experiments to empirically quantify contributions to energy conservation and latency reduction.

Funding

This research was supported by Ningbo International Investment Consulting Co., Ltd.

Disclosure statement

The author declares no conflict of interest.

References

- [1] Yang T, Kapogiannis G, Kang B, et al., 2025, A Panoramic Image-Based Framework for Building Facility

- Management: An Experimental Case Study, In Proceedings of the 2025 7th International Conference on Economic Management and Cultural Industry (ICEMCI 2025), 245–251.
- [2] Yang T, Kapogiannis G, Kang B, et al., 2025, Holographic Mixed-Reality Interfaces for Enhanced BIM-Driven Facility Management: A Case Study. *E3S Web of Conferences*, 2025(656): 01008.
- [3] Yang T, Kapogiannis G, Guo J, et al., 2025, Integrating Point Cloud in BIM for Enhancing Building Facility Management: A Case Study, *Proceedings of 7th International Conference on Civil Engineering and Architecture*, 192–207.
- [4] Becerik-Gerber B, Jazizadeh F, Li N, et al., 2012, Application Areas and Data Requirements for BIM-Enabled Facilities Management. *Journal of Construction Engineering and Management*, 138(3): 431–442.
- [5] Motawa I, Almarshad A, 2013, A Knowledge-Based BIM System for Building Maintenance. *Automation in Construction*, 2013(29): 173–182.
- [6] Gerrish T, Ruikar K, Cook M, et al., 2017, BIM Application to Building Energy Performance Visualisation and Management: Challenges and Potential. *Energy and Buildings*, 2017(144): 218–228.
- [7] Kazado D, Kavagic M, Eskicioglu R, 2019, Integrating Building Information Modeling (BIM) and Sensor Technology for Facility Management. *Journal of Information Technology in Construction*, 2019(24): 440–458.
- [8] Desogus G, Quaquero E, Rubiu G, et al., 2021, BIM and IoT Sensors Integration: A Framework for Consumption and Indoor Conditions Data Monitoring of Existing Buildings. *Sustainability*, 13(8): 4496.
- [9] Grieves M, Vickers J, 2017, Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems. *Transdisciplinary Perspectives on Complex Systems*, 85–113.
- [10] Valdiserri P, Cesari S, Coccagna M, et al., 2020, Experimental Data and Simulations of Performance and Thermal Comfort in a Patient Room Equipped with Radiant Ceiling Panels. *Buildings*, 10(12): 235.
- [11] Wang Q, Qin Q, Jiang B, et al., 2021, Mechanized Construction of Fabricated Arches for Large-Diameter Tunnels. *Automation in Construction*, 2021(124): 103583.
- [12] Deng M, Menassa C, Kamat V, 2021, From BIM to Digital Twins: A Systematic Review of the Evolution of Intelligent Building Representations in the AEC-FM Industry. *Journal of Information Technology in Construction*, 2021(26): 58–83.
- [13] Boje C, Guerriero A, Kubicki S, et al., 2020, Towards a Semantic Construction Digital Twin: Directions for Future Research. *Automation in Construction*, 2020(114): 103179.
- [14] Tao F, Zhang H, Liu A, et al., 2019, Digital Twin in Industry: State-of-the-Art. *IEEE Transactions on Industrial Informatics*, 15(4): 2405–2415.
- [15] Opoku D, Perera S, Osei-Kyei R, et al., 2021, Digital Twin Application in the Construction Industry: A Literature Review. *Journal of Building Engineering*, 2021(43): 102726.
- [16] Chen N, Lin X, Jiang H, et al., 2024, Automated Building Information Modeling Compliance Check through a Large Language Model Combined with Deep Learning and Ontology. *Buildings*, 14(7): 1983.
- [17] Zheng Z, Zhou Y, Lu X, et al., 2022, Knowledge-Informed Semantic Alignment and Rule Interpretation for Automated Compliance Checking. *Automation in Construction*, 2022(142): 104524.
- [18] Zhang J, El-Gohary N, 2017, Integrating Semantic NLP and Logic Reasoning into a Unified System for Fully-Automated Code Checking. *Automation in Construction*, 2017(73): 45–57.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.