Joint CSCE Construction Specialty & CRC Conference 2025
Conférence conjointe spécialisée en construction de la SCGC et CRC-2025

Montreal, Quebec
July 28-31, 2025 / *28-31 juillet 2025*

FROM TEXT TO MEANING: EVALUATING SEMANTIC APPROACHES FOR RULE INTERPRETATION IN AUTOMATED RULE COMPLIANCE CHECKING

E. Naderi[1]* and E. Poirier[1]

[1]Department of construction Engineering, École de Technologie Supérieure (ÉTS), Montreal, Canada

ABSTRACT:

Automated rule compliance checking (ARCC) offers an alternative to resource-intensive and error-prone manual processes. Current semantic-based approaches like ontologies, NLP, Large Language Models (LLMs) and semantic markup capture regulatory meaning, yet face challenges in accuracy, scalability, and integration. This paper evaluates and compares existing semantic-based rule interpretation methods, synthesizing lessons learned and insights from current research. Results show strengths and limitations, revealing barriers to broader adoption and opportunities for improvement. The lasting contribution is a comprehensive overview and guidance for advancing semantic-based rule preparation, aiming to enhance the effectiveness and reliability of ARCC toward fully automated, cost-effective, and error-resistant compliance checking solutions.

1. INTRODUCTION

Throughout its entire lifecycle, the built environment is guided by regulations, standards, and requirements, characterized as “rules”, which play a critical role in guaranteeing that it remains safe and fully functional (Beach et al., 2020)**.** Ensuring compliance with these rules is a complex and labor-intensive undertaking, as it remains carried out manually (Beach et al., 2020).

Moreover, while regulatory frameworks are crucial for the advancement of the built environment, the compliance checking process frequently experiences notable inefficiencies(Naderi et al., 2024). Compliance checking is the process of determining whether a solution complies with its applicable set of rules(Salama & El-Gohary, 2011). Automated rule compliance checking (ARCC) is the automated application of rules to model information about an asset, typically resulting in outcomes like pass, fail, or occasionally unknown(İlal & Altıntaş, 2022). Automated rule checking systems have a four-stage process. The first stage entails interpreting, formulating, and logically structuring rules for machine processing. In the second stage, the necessary information from built environment models is generated. The third stage—often called the rules execution phase—handles the actual rule compliance checks. Finally, in the fourth stage, the outcomes of these checks are documented and reported(Nawari, 2018).

The first step in ARCC is extracting rules from their source material, typically in natural human language, and convert them into a machine-readable format that computers can execute. Rules are commonly drafted in natural language, which computers struggle to interpret and apply with precision (Nawari, 2018).

Zhang and Ma (2022) divide the first step of ARCC process which deals with rules into three core themes: rule classification, rule organization, and interpretation and expression(Zhang & Ma, 2022).Rule interpretation and expression prepare the rules for ARCC, while rule classification and organization make this process simpler and more precise. We refer to rule interpretation and expression as rule preparation because they specifically ready the rules for ARCC.

In rule preparation, building rules are structured into a machine-readable format without altering their original meaning. Rule classification centers on categorizing individual provisions. Moreover, accounting for the dependencies between rules (i.e., rule organization) plays a crucial role, as recognizing these interconnected relationships enhances the effectiveness of rule organization and influences the outcomes of compliance checks(Zhang & Ma, 2022).

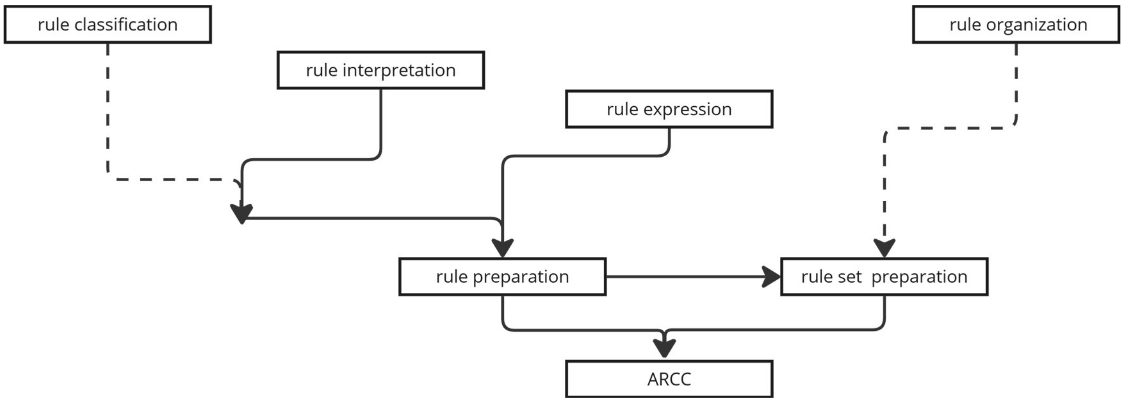

Figure 1: Main themes in the first step of ARCC process

Zheng et al. (2024) note that not all rule requirements are suitable for automated computer interpretation. Some exhibit low interpretabilities, which increases the likelihood of errors if processed automatically. Many ARCC systems encounter obstacles in rule preparation (Zhang et al., 2024).
This process can be conducted manually, semi-automatically, or fully automatically, utilizing a variety of methods—including logic-based, language-based, and semantic-based approaches (e.g., natural language processing). While the field has made considerable progress, a number of challenges still remain, like: Necessity for a procedure that transforms rules from human language into a format comprehensible by machines(Lee et al., 2016). Inability to completely automate rules compliance checking due to limitations in representing those rules in a query-able format (Hbeich et al., 2019) and the lack of a mechanism to maintain rules in an up-to-date and reusable form for various projects.
In this paper we will try to introduce and compare different semantic-based methods of rule preparation for ARCC(Lee et al., 2016).
The study is directed by the following research questions:
• What semantic approaches currently exist for preparing rules for ARCC?
• What are the benefits and limitations of each method?
• What are the prospects for rule preparation in ARCC?

A deliberate set of keywords—including “Construction Regulations,” “Built Environment Rules,” “Building Codes,” “Rule Interpretation,” “Rule Representation,” “Automated Rule Compliance Checking,” “Code Compliance,” and “Automated Rule-Checking”—guided comprehensive searches across key academic databases such as Scopus, Web of Science, and Google Scholar. The selected articles were thoroughly reviewed to identify recent advancements in automated rule compliance checking, examine novel approaches to rule interpretation and representation, and highlight notable research gaps and emerging trends. The following sections present a synthesis of these findings, illustrating current developments and outlining areas requiring further research.
Following this introduction, Section 2 outlines the pertinent research background. Section 3 summarizes key findings from the literature review, and Section 4 integrates the results and offers final conclusions, including recommendations for future investigation.

2. BACKGROUND

2.1 semantic-based representation methods

A recent focus in rule preparation research has been on semantic modeling, a comparatively new technique for knowledge representation(Macit İlal & Günaydın, 2017).This approach improves built environment regulations by incorporating semantic information elements and essential logical operators to generate individual requirement units, allowing them to be processed by machines (Zhang & El-Gohary, 2020). Mainly through a process called semantic annotation—where code provisions are assigned meaningful labels—computer systems can accurately interpret and handle these requirements (Zhong et al., 2013). A diverse set of methods developed over time for semantic rule preparation, each grounded in different technological advancements. These methods can be systematically categorized into markup methods, Natural Language Processing (NLP) methods, semantic web methods, andLarge Language Models (LLMs). (Fig .2)

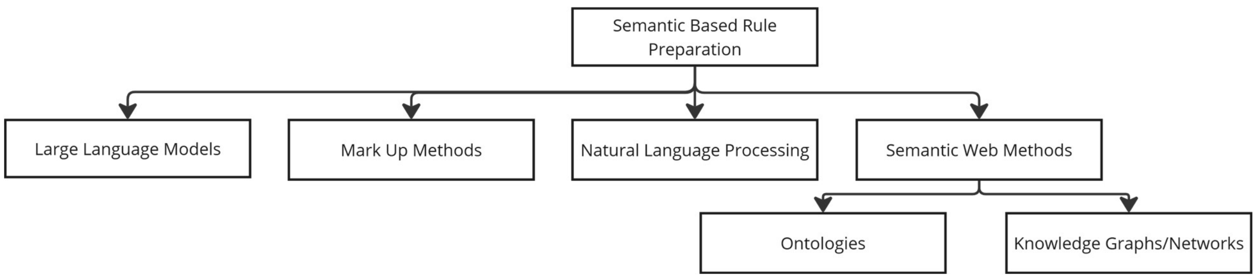

Figure 2: Semantic Based Rule Preparation methods

In the subsequent section, the findings from the evaluation of these methods are presented in a detailed, method-by-method format.

3. RESULTS

Over the few decades, efforts have been increasingly focused on transforming building regulations into automated compliance checking systems. As a result, the use of computable formats and integrating them into black-box compliance checking systems has become a common practice. This approach creates digital versions of rules that operate independently from the original documents, necessitating separate maintenance to ensure updates are synchronized. To address this, the best practices suggest closely linking the syntactic and semantic versions so that any updates to one automatically reflect in the other, which is especially beneficial for frequently revised rules and subsidiary legislation (Amor & Dimyadi, 2021). Researchers have advanced semantic-based methods to improve ARCC. For instance, a semantic analysis approach has been successfully implemented in rule-checking systems, enhancing the understanding of regulations(Song et al., 2018).Presented below are the findings from a literature review focusing on various semantic-based methods. These findings detail the characteristics of each approach, outline the associated advantages and drawbacks, and discuss the research that has been conducted in each domain.

3.1 Mark up methods

A semantic approach being explored for ARCC involves utilizing markup document modeling and hypertext to depict regulatory provisions. Numerous studies have revisited the idea of annotating regulatory texts to create machine-readable representations. (Nawari, 2018). These include RASE, XML, LegalRule ML and Dialogue Language.

3.1.1 Requirement, Applicability, Selection, and Exception (RASE)

RASE method structures rules using four semantic markup operators: Requirement (R), Applicability (A), Selection (S), and Exception (E). It retains the original regulation text while adding markups to facilitate ARCC through a tree-like representation of logical connections (Hjelseth & Nisbet, 2011; Nawari, 2018; Zhang & Ma, 2022).RASE effectively bridges the semantic gap between natural language regulations and machine-readable formats, enabling more accurate and efficient compliance checking.(Nisbet et al., 2024; Song et al., 2018; Zhang & El-Gohary, 2021).RASE requires manual identification of formal rules by domain experts and the implementation of software code based on predefined measures, which can be resource-intensive (Liu Jiang et al., 2022; Zhang & Ma, 2022).Automating the semantic parsing and markup process in RASE could reduce the dependency on manual interventions, enhancing scalability and reducing costs(Fuchs, Fauth, et al., 2024; Zheng et al., 2024).

3.1.2 Extensible Markup Language (XML)

Extensible Markup Language (XML) is extensively employed to structure the built environment regulations into formats that machines can read. It supports semi-structured data, making it well-suited for modeling legal documents (Nawari, 2018). XML allows for the creation of decision tables used in automated compliance verification and supports intelligent authoring processes for developing legislative documents. (Tan et al., 2007).XML functions as a standardized format that links diverse data types, allowing smooth data exchange across different systems. Its dual role as a markup language and a web standard increases its flexibility for use in both initial and later applications (Nawari, 2018).Even though XML successfully represents hierarchical and referential frameworks, the manual conversion of regulations into XML is both time-consuming and requires significant effort (Wyner & Governatori, 2013).

3.1.3 LegalRuleML & Dialogue Language

LegalRuleML extends XML to capture legal rules and norms with comprehensive and meaningful markup (Murat Aydın & Hakan Yaman, 2020).It combines Semantic Web technologies to model legal concepts and provisions, supporting various rule types such as facts, queries, and integrity constraints (Nawari, 2018).
LegalRuleML provides a standardized framework for representing legal rules, enhancing interoperability and enabling sophisticated automated reasoning and compliance checking (Murat Aydın & Hakan Yaman, 2020).One of the pioneering efforts in using markup languages for modeling building regulations was introduced by Omari and Roy (1993) through their Dialogue Language (DL). DL was specifically developed to interpret Life Safety Codes (LSC) for Australia within an expert system framework. It assumes a consistent interpretation of code provisions and facilitates interactions between users and the expert system. DL organizes the hierarchical patterns of regulations using eight primary structural items, providing a structured approach to representing regulatory information (Nawari, 2018).
The primary limitation of DL is its reliance on manual rule definition and updates. As regulations evolve, maintaining and updating the DL framework requires significant effort from domain experts, which can be resource-intensive and time-consuming (Nawari, 2018).

3.2 Natural Language Processing (NLP) Approaches

Several studies have adopted natural language processing (NLP), a widely used method for analyzing human language, to automate the interpretation of regulatory rules by addressing their syntactic and semantic dimensions. (Zheng et al., 2022).
AI-based language models employ various NLP algorithms, typically leveraging machine learning and deep learning techniques, to interpret and extract meaning from text.(Zheng et al., 2022).(Zheng et al., 2022).
Natural Language Processing (NLP) plays a crucial role in preparing built environment rules by automating the extraction and conversion of regulatory texts into machine-readable formats. NLP for rule preparation is defined by its capability to analyze regulatory texts as domain-specific languages, facilitating the generation of computable rules from various textual documents. This process generally involves two main steps: first, the automatic identification and extraction of semantic information from regulatory documents, known as information extraction (IE), and second, the conversion of this extracted data into logical statements that enable automated reasoning, referred to as information transformation (IT) (Xue & Zhang, 2020; Zhang & El-Gohary, 2012; Zheng et al., 2022).
Techniques such as part-of-speech tagging, tokenization, and named entity recognition are utilized to perform Information Extraction (IE) from unstructured building codes, thereby transforming them into structured, logical clauses appropriate for compliance verification (Xue & Zhang, 2020; Zhang & El-Gohary, 2012). Information transformation involves converting the extracted data into logical clauses suitable for

reasoning in compliance verification. Additionally, to enhance NLP-based methods for various rule interpretation tasks, advanced techniques like ontologies, machine learning, and deep learning have been employed to achieve a deeper and more nuanced understanding of regulatory texts(Zheng et al., 2022).
Furthermore, integrating Context-Free Grammar (CFG) with NLP enhances this functionality by offering a universal framework for parsing complex regulatory language, thereby generating precise verification rules from textual sources (Tan et al., 2007).
Furthermore, embedding techniques such as Word2Vec and BERT capture semantic relationships between words and phrases, boosting performance across various NLP tasks and enabling more accurate semantic analysis and rule extraction (Hasan et al., 2020).
Despite these advancements, NLP for rule preparation faces several challenges, including the complexity of regulatory language, the need for high-quality annotated data, and the integration of NLP systems with existing regulatory workflows(Kuppan et al., 2024; Zheng et al., 2024). Ethical and legal considerations, such as data privacy and the transparency of AI-generated outputs, also play a critical role in the deployment of NLP technologies in regulatory contexts (Chow et al., 2024).

3.3 Semantic Web Methods

Traditional web data is often too unstructured for computers to easily interpret, hindering automated information processing. (Nawari, 2018).
The Semantic Web approach is suggested to address the difficulty computers have in extracting, interpreting, and processing unstructured web data by adding contextual information (Murat Aydın & Hakan Yaman, 2020). It employs the Resource Description Framework (RDF) to represent data as a graph of logical statements, known as RDF triples, each containing a subject, predicate, and object (Akbaş, 2019).
Employing Semantic Web technologies for ARCC enables conformance to be assessed prior to construction or in its initial phases (Bernert et al., 2023).The Semantic Web approach often faces criticism for its complexity, as the steep learning curve can impede domain experts from using it effectively (Zhang & Ma, 2022).
In legal ontology research, XML annotation and Semantic Web tools address the distinct traits of legal rules and norms, leading to modeling languages such as RuleML and LegalRuleML, aimed at comprehensively representing legal provisions (Nawari, 2018).
Semantic Web approaches also involve descriptive languages, such as in Pauwels et al. (2011), who used a semantic network of directed, labeled graphs to provide acoustic regulation compliance for BIM models.

3.3.1 Ontology-Based Methods of Rule Representation

Ontologies offer an alternative means of conceptually representing rules and are vital in ARCC by providing structured, semantically rich models of both built environment and regulations (Liu Jiang et al., 2022).Ontologies are designed to capture domain-specific knowledge, including elements of the built environment, regulatory constraints, and compliance requirements(Hu et al., 2013; Zhong et al., 2012).
An ontology comprises a collection of concepts within a specific domain and the descriptions of the relationships between these concepts. Acting like an encyclopedia for computers, ontologies connect diverse metadata, clarify definitions, and list synonyms, thereby providing a unified structure for data. This enables the data to be understood and shared across multiple applications and communities (Nawari, 2018).Ontologies build knowledge bases that represent both rules and built environment information, including rule (code) ontologies, model ontologies, and merged ontologies (L. Jiang et al., 2022). A rule ontology organizes regulatory information into a structured, machine-readable format to support automated compliance checking. The model ontology captures the details of a built asset, such as a building or construction project, structured to meet regulatory requirements and streamlined through preprocessing steps to enhance the compliance checking process. Finally, the merged ontology is created by mapping the rule ontology to the model ontology, which reduces semantic ambiguities to ensure accurate compliance verification(L. Jiang et al., 2022).
An ontology-based approach improves automated compliance checking by converting regulations into machine-readable formats and integrating them with built environment models like BIM (Hu et al., 2013). Ontologies bridge the semantic gap between built environment model schemas and the natural language

of regulatory documents through ontology mapping and rule-based reasoning, which reduces ambiguities and enhances the semantic richness of built environment information (Li et al., 2023).Additionally, they enable the automatic extraction of requirements from regulatory texts using techniques such as pattern matching and dependency-based extraction, effectively managing complex and layered code provisions (Zhou & El-Gohary, 2017).Overall, the use of ontologies significantly enhances compliance automation by providing greater efficiency, faster response times, and improved data interoperability and interaction compared to traditional methods (Zhou et al., 2022).

Integrating diverse data and overcoming data silos and interoperability challenges with ontology-based solutions can present significant barriers (Farghaly et al., 2024). Additionally, merging ontologies for ARCC involves challenges related to ensuring data completeness and accuracy, as well as increasing complexity (Caselli et al., 2020). Also Ensuring interoperability and semantic consistency across data from various sources further complicates the process, requiring effective methods to combine different ontologies (Zhou et al., 2022).

Yurchyshyna et al. (2008) proposed a four-stage formal ontological approach for conformance checking, and the International Code Council's SMART codes platform created an International Energy Conservation Code dictionary for communication between rules and a building model (Akbaş, 2019; Ismail et al., 2017)Other methods use SBVR and SPARQL to redevelop building codes (Bouzidi et al., 2010) or CQIE Ontology for construction quality rule checking (Zhong et al., 2012).

3.3.2 Knowledge Graph–Driven Approaches

Built environment rules create a comprehensive and interconnected system, yet they are generally stored as plain text and PDF files, which obstructs efficient knowledge retrieval and use (Liu Jiang et al., 2022).

To tackle this issue, knowledge graphs offer a promising approach by precisely mapping the relationships within Built environment rules. These graph-based methods allow for the conversion of complex regulatory texts into computational frameworks, thereby enhancing the accessibility and functionality of the information(Liu Jiang et al., 2022; Lee et al., 2024). By employing knowledge graphs, it becomes possible to recognize building patterns through rule-based reasoning and to transform unstructured built environment rules into structured, machine-readable formats (Zhou et al., 2021).

Knowledge graphs can transform unstructured built environment rules into structured, machine-processable formats (Zhou et al., 2021). Because built environment rules reference each other, they form a vast network of knowledge, yet their storage as plain text or PDFs limits efficient retrieval and usage (Liu Jiang et al., 2022).Representing these rules as networks of nodes (objects or concepts) and edges (relationships) has become a prominent topic due to its advantages in describing and linking information (Liu Jiang et al., 2022). Consequently, knowledge graphs are viewed as a promising solution for illustrating the correlations within built environment rules(Liu Jiang et al., 2022).

Network-based methods for building regulations interpretation enhance efficiency by offering productivity improvements, deeper insights into regulatory complexities, and broader comprehension of as-built BIM for management and optimization(M. Aydın & H. Yaman, 2020; Lee et al., 2024) .They support retrieval of regulation relations, conflict analysis, and design compliance reviews (Zhou et al., 2021) while addressing challenges in translating regulations into computational constructs (Lee et al., 2024). Because built environment rules, which extensively reference each other, are often stored as plain text or PDFs, knowledge retrieval and use remain inefficient (Liu Jiang et al., 2022). As a solution, knowledge graphs provide a graph-based structure of nodes (concepts or objects) and edges (relationships), facilitating automated reasoning, rule-based checks, and more effective knowledge management (Liu Jiang et al., 2022).They can be defined as RDF graphs composed of triples (subject, predicate, object), with predicates indicating semantic relations and subjects and objects representing entities (Liu Jiang et al., 2022).

Implementing network-based methods for building regulations preparation faces challenges due to the absence of a clear methodological framework in developing representations and the complexity of interconnected elements involved in regulatory innovation (Grangaard & Lygum, 2024; Zhang et al., 2022) Zhou et al. introduced a building code graph by connecting clauses through their indexing numbers, which allows for partial retrieval but lacks semantic relationships. Meanwhile, ontologies can turn captured knowledge into interpretable, machine-readable, and explicit representations. They have been employed as meta models for construction quality inspection regulations, residential building codes, and underground utilities' spatial constraints. However, these approaches mainly focus on conceptual organization rather than detailing underlying logic, and few studies consider how to integrate built environment rule ontologies with building models across different phases of the compliance-checking process (Liu Jiang et al., 2022).

3.4 Large Language Models (LLMs)

Large Language Models (LLMs), by virtue of their substantial size and broad-ranging training datasets, are especially proficient at generating coherent text. This capability streamlines the conversion of rules into structured, semantic formats, thereby automating what has traditionally been a labor-intensive process (Fuchs, Witbrock, et al., 2024). Fuchs et al. (2024) proposed a method that capitalizes on LLMs to formalize building rules for ARCC. Because these models have been extensively pre-trained on diverse tasks, they can generate precise translations from limited examples, outperforming conventional supervised learning methods by adeptly choosing the most suitable translations for regulatory provisions—even in the absence of specialized training sets (Fuchs, Witbrock, et al., 2024).

Additionally, integrating LLMs with deep learning and ontology frameworks can further ARCC by enabling more efficient information extraction, thus reducing manual workload (Chen et al., 2024). Their capacity to bridge the semantic gap between IFC and natural language rules underscores the promising role of LLMs in aligning IFC concepts with regulatory requirements (Zhang & El-Gohary, 2021).However, handling complex legal texts with intricate nesting and conditional statements remains a significant challenge, necessitating domain-specific fine-tuning for accurate interpretation (Chen et al., 2024; Liu et al., 2023).In the Architecture, Engineering, and Construction (AEC) sector, the expansion of LLM use is additionally constrained by concerns around bias, fairness, transparency, and accountability in automated compliance processes(Liu et al., 2023).

4. DISCUSSION, CONCLUSION, FUTURE RESEARCH DIRECTIONS

4.1 Discussion

This discussion reconnects with the original research inquiries by highlighting the diverse set of methods developed over time for semantic rule preparation, each grounded in different technological advancements. These methods can be systematically categorized into markup methods, Natural Language Processing (NLP) methods, semantic web methods, and Large Language Models (LLMs), Each category leverages unique technologies and approaches to address specific aspects of rule interpretation, ensuring that the process of preparation of rules is both comprehensive and adaptable to varying regulatory requirements. For instance, markup methods excel in scenarios requiring straightforward text annotation, while ontology-based methods are ideal for environments that demand a deep understanding of contextual relationships and hierarchies. By organizing these methods into distinct categories, researchers can better understand their individual strengths and applications within the Automated Rule Compliance Checking (ARCC) process.

As mentioned in the text, each of these methods offers various benefits and is particularly suited for preparing specific types of rules based on their inherent characteristics. Researchers have continually sought to enhance and evolve their rule interpretation capabilities by combining these methods, thereby maximizing their effectiveness. Due to their complementary nature, many of these approaches naturally incorporate elements from one another, leading to more robust and versatile solutions. For example, combining NLP techniques with ontology-based frameworks can improve the accuracy of information extraction by providing contextual guidance, while integrating semantic web technologies with knowledge graphs can enhance data interoperability and reasoning capabilities. This evolutionary process has been driven by the availability of new tools and technologies, which have addressed the limitations of previous methods. Consequently, the integration and advancement of these diverse methods have significantly improved the preparation of rules for the ARCC process, ensuring greater accuracy and efficiency in regulatory compliance.

4.2 Conclusion & Future Research Directions:

This study underscores the transformative potential of semantic-based rule interpretation methodologies in advancing ARCC. By meticulously comparing and analyzing existing semantic-based approaches, including markup methods, Natural Language Processing (NLP) methods, semantic web methods, and Large Language Models (LLMs), the research highlights their capacity to convert complex regulatory rules into machine-readable formats. These methodologies offer significant advantages in terms of accuracy, scalability, and the ability to capture the underlying meaning of regulations, thereby enhancing the effectiveness and reliability of compliance checking processes.

However, the integration of semantic-based methods into existing ARCC pipelines is not without challenges. The study identifies key obstacles such as ensuring high accuracy in rule interpretation, achieving scalability across diverse regulatory frameworks, and seamlessly integrating these advanced methods with current systems. Additionally, the variability in the strengths and limitations of each approach necessitates a careful consideration of their applicability to specific regulatory contexts. To fully realize the benefits of semantic-based rule interpretation, it is imperative to address these barriers through continued research and the development of interoperable standards. Continuing research on the combination of these methods is an excellent idea. By integrating different approaches, we can leverage their strengths and create more robust solutions. Additionally, utilizing new concepts from other domains, such as computer science, can provide fresh perspectives and innovative techniques that enhance the overall effectiveness of the methods.

It is important to consider the context of the built environment in which these methods will be implemented. Understanding the specific requirements and constraints of the environment allows for the adaptation and optimization of the methods to better suit their intended applications. This careful consideration ensures that the combined and newly integrated concepts work harmoniously within the built environment, ultimately making them more efficient and effective. Moving forward, fostering collaboration among researchers, industry practitioners, and regulatory bodies will be essential to overcome these challenges and to drive the evolution of ARCC towards fully automated, cost-effective, and error-resistant compliance checking solutions.

REFERENCES

Akbaş, Y. S. (2019). *A Formal representation of building codes to facilitate BIM-based automated code-checking* Middle East Technical University].

Amor, R., & Dimyadi, J. (2021). The promise of automated compliance checking. *Developments in the Built Environment*, *5*, 100039.

Aydın, M., & Yaman, H. (2020). Domain knowledge representation languages and methods for building regulations. Communications in Computer and Information Science,

Aydın, M., & Yaman, H. (2020, 2020//). Domain Knowledge Representation Languages and Methods for Building Regulations. Advances in Building Information Modeling, Cham.

Beach, T. H., Hippolyte, J.-L., & Rezgui, Y. (2020). Towards the adoption of automated regulatory compliance checking in the built environment. *Automation in Construction*, *118*.

Bernert, M., Ramparany, F., & Hassan, T. (2023). Automating building regulations conformance checking using a semantic approach preliminary results. CEUR Workshop Proceedings,

Bouzidi, K. R., Fies, B., Bourdeau, M., Faron-Zucker, C., & Le Than, N. (2010). Towards a semantic-based approach for the creation of technical regulatory documents in Building Industry. eWork and eBusiness in Architecture, Engineering and Construction - Proceedings of the European Conference on Product and Process Modelling 2010,

Caselli, A., Daponte, V., Falquet, G., & Métral, C. (2020). A rule language model for subsurface data refinement. EG-ICE 2020 Workshop on Intelligent Computing in Engineering, Proceedings, Chen, N., Lin, X., Jiang, H., & An, Y. (2024). Automated Building Information Modeling Compliance Check through a Large Language Model Combined with Deep Learning and Ontology [Article]. *Buildings*, *14*(7), Article 1983.

Chow, J. C. L., Wong, V., & Li, K. (2024). Generative Pre-Trained Transformer-Empowered Healthcare Conversations: Current Trends, Challenges, and Future Directions in Large Language Model-Enabled Medical Chatbots [Review]. *BioMedInformatics*, *4*(1), 837-852.

Farghaly, K., Soman, R., & Whyte, J. (2024). cSite ontology for production control of construction sites [Article]. *Automation in Construction*, *158*, Article 105224.

Fuchs, S., Fauth, J., & Amor, R. (2024). Developing Formal Representations of Building Regulations through a Parallel Drafting Process. Proceedings of the 31st International Workshop on Intelligent Computing in Engineering, EG-ICE 2024,

Fuchs, S., Witbrock, M., Dimyadi, J., & Amor, R. (2024). Using Large Language Models for the Interpretation of Building Regulations [Article]. *Journal of Engineering, Project, and Production Management*, *14*(4), Article 0035.

Grangaard, S., & Lygum, V. L. (2024). Unveiling the missing innovation – an actor–network theory analysis of regulatory efforts towards universal design [Article]. *Archnet-IJAR: International Journal of Architectural Research*.

Hasan, F., Roy, A., & Pan, S. (2020). Integrating Text Embedding with Traditional NLP Features for Clinical Relation Extraction. Proceedings - International Conference on Tools with Artificial Intelligence, ICTAI,

Hbeich, E., Roxin, A., & Bus, N. (2019). Applying PLU Rules on Different BIM/GIS Interoperability Approaches. *Informatica Economica*, *23*(3/2019), 26-38.

Hjelseth, E., & Nisbet, N. N. (2011). Capturing normative constraints by use of the semantic mark-up RASE methodology.

Hu, Y. Z., Zhong, B., Luo, H., & Hu, H. M. (2013). The regulation constraint modeling and semantic inferring in construction quality checking [Article]. *Structural Survey*, *31*(5), 387-400.

İlal, M. E., & Altıntaş, Y. D. (2022). Integrating building and context information for automated zoning code checking: a review. *Journal of Information Technology in Construction*, *27*, 548-570.

Ismail, A. S., Ali, K. N., & Iahad, N. A. (2017). A Review on BIM-based automated code compliance checking system. International Conference on Research and Innovation in Information Systems, ICRIIS,

Jiang, L., Shi, J., Pan, Z., Wang, C., & Mulatibieke, N. (2022). A Multiscale Modelling Approach to Support Knowledge Representation of Building Codes. *Buildings*, *12*(10), 1638.

Jiang, L., Shi, J., & Wang, C. (2022). Multi-ontology fusion and rule development to facilitate automated code compliance checking using BIM and rule-based reasoning [Article]. *Advanced Engineering Informatics*, *51*, Article 101449.

Kuppan, K., Acharya, D. B., & Divya, B. (2024). Foundational AI in Insurance and Real Estate: A Survey of Applications, Challenges, and Future Directions [Article]. *IEEE Access*.

Lee, C., Yeo, K., & Dritsas, S. (2024). BIM-ENABLED REGULATORY DESIGN RULE CHECKING FOR BUILDING CIRCULATION. Proceedings of the International Conference on Computer-Aided Architectural Design Research in Asia,

Lee, H., Lee, J.-K., Park, S., & Kim, I. (2016). Translating building legislation into a computer-executable format for evaluating building permit requirements. *Automation in Construction*, *71*, 49-61.

Li, Y., Yang, M., Zhao, Q., Li, Z., Ma, Z., Liu, Y., & Hei, X. (2023). A Semantic Representation Method of Building Codes Applied to Compliance Checking [Article]. *Mathematics*, *11*(11), Article 2552.

Liu, X., Li, H., & Zhu, X. (2023). A GPT-based method of Automated Compliance Checking through prompt engineering. 30th EG-ICE: International Conference on Intelligent Computing in Engineering 2023,

Macit İlal, S., & Günaydın, H. M. (2017). Computer representation of building codes for automated compliance checking. *Automation in Construction*, *82*, 43-58.

Naderi, E., Poirier, E., & Jalaei , F. (2024). From Paperwork to the Frontiers of Knowledge and Technology: What Role Can the Integration of BIM and GIS Play in Regulatory Compliance? Proceedings of the 41st International Conference of CIB W78, Marrakech, Morocco.

Nawari, N. O. (2018). *Building Information Modeling: Automated Code Checking and Compliance Processes* CRC Press.

Nisbet, N., Zhang, Z., & Çidik, S. (2024). REAL-TIME ASSESSMENT OF REGULATORY COMPLIANCE OF CONSTRUCTION SITES. Proceedings of the European Conference on Computing in Construction,

Salama, D., & El-Gohary, N. (2011). *Semantic Modeling for Automated Compliance Checking*.

Song, J., Kim, J., & Lee, J. K. (2018). NLP and deep learning-based analysis of building regulations to support automated rule checking system. ISARC 2018 - 35th International Symposium on Automation and Robotics in Construction and International AEC/FM Hackathon: The Future of Building Things,

Tan, X., Hammad, A., & Fazio, P. (2007). Automated code compliance checking of building envelope performance. Congress on Computing in Civil Engineering, Proceedings,

Wyner, A., & Governatori, G. (2013). A study on translating regulatory rules from natural language to defeasible logic. CEUR Workshop Proceedings,

Xue, X., & Zhang, J. (2020). Building Codes Part-of-Speech Tagging Performance Improvement by Error-Driven Transformational Rules [Article]. *Journal of Computing in Civil Engineering*, *34*(5), Article 04020035.

Zhang, J., & El-Gohary, N. (2012). Automated regulatory information extraction from building codes leveraging syntactic and semantic information. Construction Research Congress 2012: Construction Challenges in a Flat World, Proceedings of the 2012 Construction Research Congress,

Zhang, J., Yu, Z., & Zhao, L. (2024). Representation of structural design specifications based on first-order predicate logic [Article]. *Tumu yu Huanjing Gongcheng Xuebao/Journal of Civil and Environmental Engineering*, *46*(1), 254-262.

Zhang, R., & El-Gohary, N. (2020). A Machine-Learning Approach for Semantic Matching of Building Codes and Building Information Models (BIMs) for Supporting Automated Code Checking. Sustainable Civil Infrastructures,

Zhang, R., & El-Gohary, N. (2021). Semantic Representation Learning and Information Integration of BIM and Regulations. Computing in Civil Engineering 2021 - Selected Papers from the ASCE International Conference on Computing in Civil Engineering 2021,

Zhang, Z., & Ma, L. (2022). *Towards fully-automated code compliance checking of building regulations: challenges for rule interpretation and representation*.

Zhang, Z., Ma, L., & Broyd, T. (2022). TOWARDS FULLY-AUTOMATED CODE COMPLIANCE CHECKING OF BUILDING REGULATIONS: CHALLENGES FOR RULE INTERPRETATION AND REPRESENTATION. Proceedings of the European Conference on Computing in Construction,

Zheng, Z., Zhou, Y.-C., Lu, X.-Z., & Lin, J.-R. (2022). Knowledge-informed semantic alignment and rule interpretation for automated compliance checking. *Automation in Construction*, *142*, 104524.

Zheng, Z., Zhou, Y. C., Chen, K. Y., Lu, X. Z., She, Z. T., & Lin, J. R. (2024). A text classification-based approach for evaluating and enhancing the machine interpretability of building codes [Article]. *Engineering Applications of Artificial Intelligence*, *127*, Article 107207.

Zhong, B., Hu, Y., & Luo, H. (2013). Study on building code ontological modeling and construction quality checking application [Article]. *Tumu Gongcheng Xuebao/China Civil Engineering Journal*, *46*(8), 136-142.

Zhong, B. T., Ding, L. Y., Luo, H. B., Zhou, Y., Hu, Y. Z., & Hu, H. M. (2012). Ontology-based semantic modeling of regulation constraint for automated construction quality compliance checking [Article]. *Automation in Construction*, *28*, 58-70.

Zhou, P., & El-Gohary, N. (2017). Ontology-based automated information extraction from building energy conservation codes [Article]. *Automation in Construction*, *74*, 103-117.

Zhou, Y., Lin, J., & She, Z. (2021). Automatic Construction of Building Code Graph for Regulation Intelligence. ICCREM 2021: Challenges of the Construction Industry under the Pandemic - Proceedings of the International Conference on Construction and Real Estate Management 2021,

Zhou, Y., She, J., Huang, Y., Li, L., Zhang, L., & Zhang, J. (2022). A Design for Safety (DFS) Semantic Framework Development Based on Natural Language Processing (NLP) for Automated Compliance Checking Using BIM: The Case of China [Article]. *Buildings*, *12*(6), Article 780.