AI-BASED DIGITAL METHODS FOR CONSTRUCTION DOCUMENTATION EXAMINATION

Zafar Matniyazov
Zilola Rakhmatillaeva
Samidullo Elmurodov
Nizomiddin Buronov
Tajibayev Jurat
[1]Tashkent University of Architecture and Civil Engineering
[2]Alfraganus University

ABSTRACT*: The article reviews modern approaches to automating the regulatory examination of architectural and construction documentation using artificial intelligence technologies. The study analyzes the primary tasks involved in the examination of design solutions, specifically the verification of compliance between design parameters and building codes and requirements. Methods for processing textual and graphical information are described, including NLP (Natural Language Processing), computer vision (technical drawing analysis), formal ontologies and knowledge graphs, as well as machine learning and Large Language Models (LLMs). The paper provides an overview of existing solutions and systems that utilize the classification of regulatory documents [1], neural network-based identification of BIM model elements [2], and contextual semantic analysis and rule generation [3]. A comparative analysis of these approaches is performed, identifying their strengths and weaknesses within the context of expert examination. The study discusses current challenges (diversity and complexity of regulations, source data quality, and the need for adaptation to the Russian language) and future prospects (integration of LLMs with BIM, increased digitalization, and data standardization). It is demonstrated that the implementation of AI solutions significantly accelerates project verification and reduces the volume of routine operations for experts [4]; however, it requires further development of methodologies for verification and ensuring the reliability of conclusions. The presented work serves as a reference for researchers and practitioners in the development and implementation of digital systems for design documentation examination.*

KEYWORDS*: artificial intelligence, regulatory examination, construction documentation, automated compliance checking, computer vision, natural language processing, ontologies, large language models.*

Latin American Journal of Education

INTRODUCTION

The examination of design documentation is a critically important stage in ensuring the safety and compliance of construction objects with regulations. Traditionally, the verification of design solutions is conducted by specialists manually analyzing drawings, cost estimates, and textual documents for compliance with SNiP, SP, GOST, and other regulations. Such an approach is labor-intensive and error-prone, especially when processing large volumes of data [4][2]. In recent years, active digitalization of the examination process has been observed: electronic document management systems and databases of estimates and conclusions are being created. Glavgosexpertiza of Russia has developed and is implementing predictive analytics and AI modules, indicating a trend towards automated verification methods. In 2023–2024, state structures began the transition of all regulatory documents into a machine-readable format [4], and the professional expert community is developing a "roadmap" for the application of neural networks in examination processes.

The role of AI lies in automating routine operations (extracting key parameters, checking basic compliances) and providing experts with tools for the deep analysis of complex cases. In particular, AI technologies allow for the processing of regulatory texts and project drawings in digital form, identifying hidden dependencies, and offering recommendations. Initially applied in medicine and education, machine learning methods and Large Language Models (LLMs) are now viewed as "smart assistants" in construction examination [4].

Thus, the objective of this work is the systematization and critical analysis of modern digital methods for construction documentation examination, with a focus on the application of AI in checking project compliance with regulations. The article is structured as follows. First, the theoretical foundations of information representation and processing (regulations and design data) are described. Then, an overview of existing systems and research results demonstrating various approaches (NLP, computer vision, ontologies, LLM, etc.) is provided. A comparative analysis of these methods regarding their capabilities and limitations is performed. The "Challenges and Prospects" section discusses the main problems and expected directions of technology development. The conclusion summarizes the findings and outlines promising topics for further research.

Theoretical Foundations

Design documentation examination is based on the analysis of several types of data:

- Textual regulatory documents (SNiP, SP, GOST, Technical Regulations, etc.) containing descriptive requirements (premises area, loads, fire resistance, etc.). These texts often formulate requirements in natural language and

under complex conditions that are sometimes difficult to formalize.

- Project texts (structural descriptions, explanatory notes, cost estimates) containing numerical data and object parameters.

- Graphical materials (drawings, floor plans, facades, BIM models) containing geometric and visual information.

The primary goal is to establish whether the design documentation parameters comply with regulatory requirements. This necessitates compliance rules (extracted from regulations) and methods for applying them to the design solution data.

Key AI technologies addressing this task can be categorized into several groups:

- **Natural Language Processing (NLP).** NLP is applied for the automatic extraction of regulations from text, including the recognition, classification, and formalization of requirements. NLP methods allow for the segmentation of regulatory texts into individual clauses and the determination of their semantic roles (e.g., "minimum room height"). This information is subsequently used to locate corresponding elements within the project. NLP tasks encompass tokenization, syntactic analysis, and the construction of semantic models of regulations. In particular, algorithms have been developed for the semantic classification of text fragments within regulatory documents regarding requirement categories (occupational

safety, fire safety, etc.) [1][5]. Successful examples demonstrate that AI classifiers can identify document sections relevant to the domain under review with high precision (for instance, Salama et al. (2016) achieved ~96% accuracy in clause classification [1]).

- **Computer Vision (CV).** CV is utilized for the automated interpretation of technical drawings and BIM models. This involves the recognition of objects (doors, windows, stairs) and their parameters in images or BIM data using neural networks. Specific tasks include building geometry extraction, layout element identification, and the measurement of distances and areas. For example, Convolutional Neural Networks (CNNs) can be used to verify object accessibility: Lange et al. (2021) applied a convolutional network to BIM models for the automated verification of design compliance with accessibility requirements, achieving ~95% accuracy [2]. Such methods allow, for example, for the automatic determination of ceiling heights or the presence of fire exits on a plan, followed by comparison with regulatory limits.

- **Ontologies and Knowledge Graphs.** The ontological approach involves the formal modeling of the examination domain: the project and regulations are described as structured graphs (concepts, properties, relations). For instance, OntoBPR (Hagedorn et al., 2025)—an ontology for the building permit issuance process—links project entities (building, floor, structural

element) with regulations [3]. A semantic model (knowledge graph) enables logical inference: for example, if a room's height is below the regulatory minimum, the system flags a violation. Ontologies also facilitate the integration of heterogeneous information (textual rules, BIM objects). Combining ontologies with semantic validation technologies (SHACL validation) allows for the automated verification of project data consistency with prescribed constraints [3].

- **Machine Learning and Deep Learning.** Traditional ML methods (SVM, decision trees) and deep neural networks are employed to identify patterns and anomalies. For example, training on historical verification data can reveal typical design errors. Neural networks can predict non-compliance categories based on a complex set of BIM model features. Furthermore, Large Language Models (LLMs) (ChatGPT, GPT-4, Gemini, etc.) offer new capabilities: they are able to answer questions based on regulatory texts, generate verification scripts, and identify hidden connections within documents. In systems integrated with BIM, LLMs can interpret complex regulations, transforming them into programmable rules. An example is a modern system integrating ChatGPT with Revit software for semi-automated verification; such an architecture has demonstrated an acceleration of the examination process and improved accuracy in violation detection [2].

Fig. 1 presents a general schematic of a simplified neural network, one of the key machine learning tools used for analyzing design data. The illustration demonstrates the concept of signal transmission between nodes (neurons), analogous to the process of pattern recognition in drawings or semantic extraction from text. Modern solutions for design
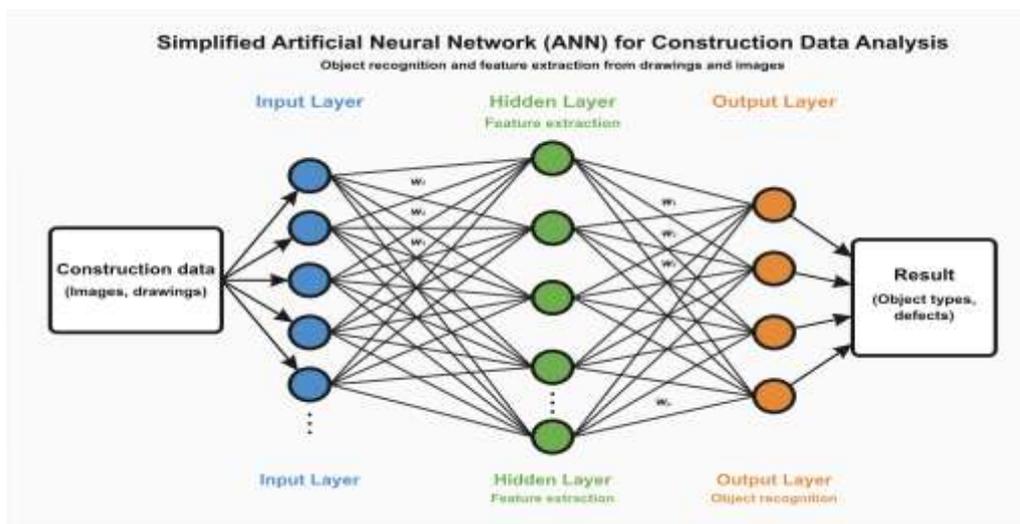
learning),

Fig. 1. Simplified schematic of an Artificial Neural Network (ANN) – a key component of deep learning methods used for construction data analysis (object recognition and feature extraction)

documentation examination often employ complex network structures (deep comprising hundreds of such neurons and layers, which allows for the effective

identification of patterns and the assessment of regulatory compliance.

Overview of Existing Solutions

In recent years, numerous initiatives and systems have emerged in the field of automated verification of construction projects. This section examines the most significant among them, providing examples of both international and domestic developments.

- Classification and Requirement Extraction from Text. Salama and El-Gohary (2016) proposed an NLP approach for the automated classification of text fragments within general contract conditions [1], serving as a foundation for subsequent verification. Their system semantically classifies clauses regarding environmental safety, health, etc., achieving high accuracy. Similarly, Li et al. (2024) utilize a combination of NLP and ontologies to extract regulatory rules within the Chinese construction market [2]. Methods for parsing Russian regulations are also being developed in Russia: experiments on preprocessing Russian text of urban regulations have demonstrated the feasibility of parameter extraction using the NLTK library (Matos Leite de Ávila et al., 2024) [1]. Thus, NLP models allow for the formalization of a portion of regulatory requirements for subsequent automated verification.

- Automated BIM Model Checking. BIM technology provides an integrated representation of design data. ACC (Automated Compliance Checking) systems compare the BIM model against formalized rules. For instance, Chen et al.

(2024) employ LLMs in conjunction with ontologies for checking BIM models [2]. Madireddy et al. (2025) developed a platform integrating various LLMs (ChatGPT, Claude, Llama) with Revit; it interprets building codes, generates Python scripts, and performs semi-automated verification, which accelerates analysis and reduces the number of omissions [2]. These studies indicate that the combination of BIM and LLMs facilitates the generation of adaptive rules and improves the reliability of violation detection.

- Cross-Language and Multimodal Approaches. Nakhaee et al. (2024) described a concept combining Knowledge Graphs (KGs) and LLMs [5]. Their hybrid approach involves integrating construction data and regulations into a unified KG, while the LLM is used for navigation and inference within this structure. Such a scheme allows for a more comprehensive consideration of context and logical connections during verification. The study demonstrates that a hybrid KG+LLM model can significantly enhance system adaptability amidst changing regulations.

- Ontologies and Expert Systems. Research indicates the effectiveness of ontology-based expert systems. For example, OntoBPR (Hagedorn et al., 2025) formalizes a multi-level ontology for the digitalization of project approval [3]. Another project, OBPA, creates a quality ontology (Ontology for Building Quality Assurance), used for inspections

Latin American Journal of Education

and control. These systems enable the formalization of complex regulatory constructs and their automated compliance verification. In automated examination management projects, ontologies serve to verify design templates and issue recommendations.

Overall, the review shows that AI tools are already being implemented in many aspects of examination. Traditional methods (rules, mathematics) are being supplemented by machine learning, semantic, and hybrid approaches. In practice, ready-made software products are used (e.g., automated BIM checking systems from Bentley, NormaCS, etc.), although they are predominantly oriented towards foreign regulations. Academic developments demonstrate how technologies are translated from research into solutions.

Comparative Analysis of Methods

All reviewed methods possess specific strengths and weaknesses when applied to the examination of design documentation. Table 1 presents the main categories of AI approaches, their areas of application, as well as their characteristic features and limitations.

Table 1. Brief characterization of AI methods in design documentation examination.

I Method	Application Area	Examples/Notes	Advantages	Limitations
NLP (Natural Language Processing)	Extraction and formalization of requirements from textual regulations and project descriptions.	Classification of documentation fragments by category (security, safety); natural language analysis [1][3].	Allows processing large volumes of regulations, automatically structuring and indexing rules.	Difficult to process vague formulations and context-dependent conditions; requires careful tokenization and parsing.
Computer Vision (CV)	Analysis of graphical materials: drawings, plans, BIM elements.	Object recognition on plans (windows, doors), assessment of geometric parameters; e.g., accessibility checking via CNN [2].	Automates the analysis of complex drawings and BIM models, identifies geometric non-compliance with requirements.	Requires large labeled datasets; sensitive to image quality and drawing standardization.
Ontol	Formaliza	Ontological	Ensures	Developmen

ogies / Knowledge Graphs	tion of relationships between project objects and regulations.	examination models (OntoBPR) for project approval [3]; rules in SHACL format.	strict formalism, capable of processing complex logical connections and automating inference.	t and maintenance of ontologies is complex, requires expert domain knowledge; dependent on model completeness.
Machine Learning / LLM	Training on historical data, error prediction, generation of textual explanations.	Violation classification models; Large Language Models (ChatGPT, GPT-4) for answering regulation-related questions [2][5].	Flexibility and adaptability to complex data patterns; ability to self-learn on new examples.	High computational costs; potential for "hallucinations" in LLMs; difficult to ensure full control over output.

As evident from the table, the choice of approach depends on the specific examination task. NLP and semantic methods are well-suited for understanding textual rules but require rich linguistic resources and often fail to account for the context of technical drawings. Computer Vision (CV) algorithms are effective in processing graphical data; however, they necessitate large volumes of labeled examples and strict drawing standards. Ontologies enable complete formalism and the integration of heterogeneous information, but their creation is difficult to scale. Machine learning, particularly large models, offers flexibility and savings in manual effort, yet requires powerful servers and may produce inaccurate results without proper verification.

The literature notes that a combination of approaches often yields the optimal result. For instance, integrating a semantic model with an LLM (Nakhaee et al., 2024 [5]) allows for leveraging the advantages of both technologies: the LLM assists in interpreting complex requirements, while the knowledge graph ensures the consistency and explainability of the decision.

CHALLENGES AND PROSPECTS

Despite the successes achieved, the automated examination of design documentation faces a number of problems.

First, the diversity and complexity of regulations. Building codes vary regionally and temporally; textual formulations are far from formal logical

constructions. Continuous updating of language models is necessary: the system must "know" new SNiPs and GOSTs.

Second, the quality of source data. Design documents often contain errors and unformatted text. As noted in industry reports, a massive volume of unstructured documents has accumulated over the years, the manual processing of which is "practically impossible" without AI [4]. Here, AI technologies themselves depend on the quality of training data, leading to the "garbage-in, garbage-out" problem. The ambiguous verification of model conclusions (especially regarding LLMs) presents another difficulty: how to ensure that the automated system has correctly interpreted the regulation and has not committed an error?

Third, the integration of AI into existing processes. Examination is a procedure with strict regulations and liability requirements. Therefore, AI solutions must be transparent and compliant with standards. It is also necessary to train experts to work with AI modules: as noted by I. Manylov, the goal is to "transfer all routine functions to machines, while the expert remains for creative tasks" [4]. This approach requires personnel training and methodological documents regulating work with AI.

Future development prospects are linked to the further growth of construction digitalization. BIM and unified information models open new opportunities: the standardization of exchange formats (e.g., IFC) will allow for broader application of AI systems.

Large Language Models are evolving rapidly: their ability to learn from voluminous text corpora makes the creation of a "digital expert" capable of understanding regulations at a specialist level more realistic. Today, LLMs already demonstrate potential in tasks involving the derivation of regulations from complex texts [2].

The development of hybrid systems—combining classical programming and AI—is also promising. For example, a system can automatically generate verification scripts (as in Madireddy et al., 2025 [2]) but require confirmation of critical decisions based on stricter rules. There is increasing interest in contrastive learning and symbolic methods that combine machine learning with clear rules. In technical regulation, an infrastructure of "smart" standards—regulations in a machine-readable format—is being built, which will facilitate AI implementation.

Thus, while obstacles exist on the path to implementing digital examination, they are gradually being successfully overcome. The expert knowledge base and accumulated examination data serve as "fuel" for AI: training datasets already contain tens of millions of text fragments from conclusions [4]. This opens prospects for creating increasingly accurate verification algorithms. The multiplicative effect of applying AI in examination can extend to the entire construction industry [4]: improved project quality and reduced verification times will lead to a more efficient

Latin American Journal of Education

investment process.

CONCLUSIONS

Modern practice in the examination of design documentation stands on the verge of a digital transformation. The application of AI methods (NLP, computer vision, ontological inference, machine learning, and Large Language Models) already provides significant acceleration of project verification and a reduction in the volume of routine operations [4].

The key findings of the review include:

- Ontologies and semantic models demonstrate high flexibility and formal rigor in automating regulatory codes; however, they require significant development effort.

- Trainable models and LLMs ensure adaptability and handle unstructured data (text, plans), but necessitate powerful computational resources and high-quality training datasets.

-

- Computer vision effectively recognizes objects in technical drawings and BIM models but remains sensitive to format diversity and image quality.

- Government initiatives confirm the practical necessity of AI: the digitalization of regulations and the implementation of pilot AI modules are already underway [4].

Thus, a combination of technologies appears to be the most fruitful approach: for example, LLMs can automatically formulate verification conditions based on ontologies and BIM data [5][2]. The transition to the widespread use of such systems will require data unification, format standardization, and sound regulatory policy. Overall, the further development of digital examination will provide experts with higher-quality tools, while offering design organizations increased speed and reliability in the examination process.

REFERENCES:

1. Salama, D. M., & El-Gohary, N. M. (2016). Semantic text classification for supporting automated compliance checking in construction. *Journal of Computing in Civil Engineering, 30*(1), 04014106[1]. https://doi.org/10.1061/(ASCE)CP.1943-5487.0000301

2. Madireddy, S., Gao, L., Ud Din, Z., Kim, K., Senouci, A., Han, Z., & Zhang, Y. (2025). Large language model-driven code compliance checking in Building Information Modeling. *Electronics, 14*(11), 2146[11]. https://doi.org/10.3390/electronics14112146

3. Hagedorn, P., Fauth, J., Zentgraf, S., Seiß, S., König, M., & Brilakis, I. (2025). OntoBPR: An ontology-based framework for performing building permit reviews using standardized information containers. *Advanced Engineering Informatics, 61*, 103369[3]. https://doi.org/10.1016/j.aei.2025.103369

4. ФАУ «Главгосэкспертиза России». (2023, December 4). Искусственный интеллект «зашел» на экспертизу. https://gge.ru/press-center/news/iskusstvennyy-intellekt-zashel-na-ekspertizu/

5. Nakhaee, A., Elshani, D., & Wortmann, T. (2024). A vision for automated building code compliance checking by unifying hybrid knowledge graphs and large language models. In P. Eversmann (Ed.), *Scalable Disruptors (Design Modelling Symposium 2024)* (pp. 445–457)[13]. Springer. https://doi.org/10.1007/978-3-031-68275-9_36

6. Fuchs, S., Witbrock, M., Dimyadi, J., & Amor, R. (2024). Using large language models for the interpretation of building regulations. *ArXiv preprint* arXiv:2407.21060.

7. Buronov, N. S., Rakhmatillaeva, Z., Matniyazov, Z., Arabi, F., & Husainov, M. (2025). *Advancing the understanding and application of building information modeling.* American Journal of Education and Learning, 3(3), 998–1006. https://doi.org/10.5281/zenodo.15083900

8. Matniyazov, Z., Giyosov, I., Rakhmatillaeva, Z., Buronov, N., & Nigmadjanova, A. (2025). *Requirements for the preparation of design documentation based on BIM technology.* American Journal of Education and Learning, 3(3), 985–991. https://doi.org/10.5281/zenodo.15083815

9. Zilola Rakhmatillaeva. (2025). Spatial Analysis of Contemporary One-Room Apartment Layouts in Tashkent's New Residential Buildings. International Multidisciplinary Journal for Research & Development, 12(07). Retrieved from https://www.ijmrd.in/index.php/imjrd/article/view/3565

10. Rakhmatillaeva, Z. *Design generation based on artificial intelligence: A comparative analysis of methodological approaches.* Presented at the International Conference: Innovations in Science and Education System, Dehli, India. 2025. https://eijmr.org/conferences/index.php/eimrc/article/view/1179

11. Rakhmatillaeva, Z. *Evolution of artificial intelligence and its integration in architectural design.* Presented at the International Conference: Innovations in Science and Education System, Dehli, India. 2025. https://eijmr.org/conferences/index.php/eimrc/article/view/1180/1411

12. Rakhmatillaeva Z.Z., Matniyazov Z.E. *Integration of artificial intelligence technologies into the landscape design process: roles, benefits, and limitations across design stages* // Collection of Research Papers. LinguaConnect: Global Perspectives on Modern Language Education. – Tashkent: Worldly Knowledge Publishing Centre, 2025. – P. 103–107. – Available at: https://www.wosjournals.com/index.php/ruconf/article/view/3303

13. Rakhmatillaeva Z.Z., Matniyazov Z.E. *AI and immersive technologies in architectural design education* // Collection of Research Papers. LinguaConnect: Global Perspectives on Modern Language Education. – Tashkent: Worldly Knowledge Publishing

Latin American Journal of Education

Centre, 2025. – P. 108–109. – Available at: https://www.wosjournals.com/index.php/ruconf/article/view/3304

14. Raxmatillayeva, Z. *ChatGPT yordamida bir xonali kvartiraning rejasini vizual tahlil qilish va optimallashtirish tajribasi.* Ilm-fan yangiliklari konferensiyasi, Andijon. 226–242-b. 2025, July. https://worldlyjournals.com/index.php/ztvdq/article/view/14115/18272

15. Matniyazov, Z. (2025). *The role and potential of BIM in digital design. American Journal of Education and Learning, 3(7), 151–169.* https://doi.org/10.5281/zenodo.16139023

16. Matniyazov, Z. (2025). *Digital transformation of the building lifecycle. American Journal of Education and Learning, 3(7), 171–189.* https://doi.org/10.5281/zenodo.16139207

17. Abdullaev, U., Dzhusuev, U., Asanova, S., Matniyazov, Z., & Pavlovskyi, S. (2025). *Research into modern methods of producing energy-efficient building materials. Architecture Image Studies, 6(1), Territories.*

18. **Raxmatillayeva, Z.** *Sun'iy intellekt asosida yaratilgan render va professional vizualizatsiya o'rtasidagi qiyosiy tahlil. Ilm fan yangiliklari konferensiyasi*, Andijon, Uzbekistan, 119–123. 2025, July. https://worldlyjournals.com/index.php/ztvdq/article/view/13966/18125

19. Matniyazov, Z. E., & Bo'ronov, N. S. (2025). AN'ANAVIY VA BIM LOYIHALASH TEXNOLOGIYALARI INTEGRATSIYASI. *Xalqaro ilmiy-amaliy konferensiyalar, 1(4),* 87–109. https://innoworld.net/index.php/ispconference/article/view/787

20. Matniyazov, Z., Tulaganov, B., Adilov, Z., Khadjaev, R., & Elmurodov, S. (2025). *Application of BIM technologies in building operating organizations.* American Journal of Education and Learning, 3(3), 957–964. https://doi.org/10.5281/zenodo.15081913

21. Matniyazov, Z., Adilov, Z., Khotamov, A., Elmurodov, S., Rasul-Zade, L., & Abdikhalilov, F. (2025). *Integration of BIM and GIS technologies in modern urban planning: Challenges and prospects.* American Journal of Education and Learning, 3(3), 972–976. https://doi.org/10.5281/zenodo.15083760

22. Isroilova, N. F., Matniyazov, Z. E., & Mansurov, Y. M. (2022). Modern trends in interior design of hotel premises. Eurasian Journal of Engineering and Technology, 5, 55-59.

23. Quldosheva, R. U., Matniyazov, Z. E., & Mansurov, Y. M. (2022). Technological Equipment of Modern Kitchen. Eurasian Journal of Engineering and Technology, 5, 28-32.

24.   Matniyazov, Z. E. (2020). Cultural and cognitive aspect and factors influencing the organization of the architectural environment of the aralsea region tourist routes. PalArch's Journal of Archaeology of Egypt/Egyptology, 17(6), 8139-8153.