

A SYSTEMATIC LITERATURE REVIEW ON THE INTEGRATION OF ARTIFICIAL INTELLIGENCE IN INFORMATION SYSTEM REQUIREMENTS ANALYSIS

Risqi Akbar Matsuka¹, Prayogo Bagus Sudarmaji^{*2}, Zain Nur Zaman³, Ilham Albana⁴

^{1,2,3,4}Information Technology, Faculty of Computer Science, Universitas Amikom, Purwokerto, Indonesia

Email: 1matsukarisqiakbar@gmail.com, 2prabaskara@gmail.com, 3zainnurzaman333@gmail.com,

4ilhamalbana@amikompurwokerto.ac.id

(Received: April 20, 2026; Revised: April 26, 2026; Published: May 4, 2026)

Abstract

Requirements analysis is a critical phase in the development of information systems, as it significantly influences the overall success of a system. However, traditional approaches to requirements analysis are often performed manually and are prone to errors, inconsistencies, and inefficiencies. The advancement of Artificial Intelligence (AI) provides new opportunities to improve the effectiveness and automation of this process. This study aims to analyze the integration of AI in requirements analysis using a Systematic Literature Review (SLR) approach. The review follows the PRISMA 2020 guidelines and examines relevant studies published between 2020 and 2025. A total of 14 selected articles were analyzed to identify commonly used AI techniques, evaluate their effectiveness, and explore existing challenges. The results indicate that various AI techniques, including Machine Learning, Deep Learning, Transformer-based models, and Large Language Models (LLMs), have been widely applied in requirements analysis tasks such as classification, ambiguity detection, information extraction, and prioritization. These techniques demonstrate improvements in accuracy, time efficiency, and consistency compared to conventional methods. Despite these advantages, several challenges remain, including data imbalance, limited model generalization, lack of explainability, and limited validation in real-world industrial environments. Therefore, further research is needed to enhance the reliability and applicability of AI-based approaches in practical settings.

Keywords: artificial intelligence; requirements engineering; systematic literature review; transformer; machine learning.

1. INTRODUCTION

Requirements analysis is a fundamental stage in the development of information systems, playing a crucial role in determining the success of a system. At this stage, user requirements are identified, analyzed, and documented as the foundation for the software development process. Errors in requirements analysis can significantly impact development costs, cause project delays, and even lead to overall system failure. Therefore, the quality of the requirements analysis process is a critical factor in producing systems that meet user needs [1], [2], [3]. However, traditional requirements analysis processes are still largely performed manually through interviews, observations, and analysis of unstructured textual documents. This approach has several limitations, including susceptibility to misinterpretation, inconsistency among analysts, and relatively long processing times [4], [5]. In addition, the increasing complexity of information systems and the growing volume of requirement data further intensify the challenges in ensuring accuracy and efficiency in the requirements analysis process. The advancement of Artificial Intelligence (AI) technologies provides new opportunities to address these challenges. Various AI techniques, such as Machine Learning, Natural Language Processing (NLP), Deep Learning, as well as Transformer-based models and Large Language Models (LLMs), have been widely utilized to support the automation of requirements analysis processes [6], [7], [8]. The application of AI enables tasks such as requirements classification (functional and non-functional requirements), ambiguity detection, information extraction from unstructured documents, and requirements prioritization to be performed more efficiently and accurately [9], [10].

Several studies have shown that the use of AI in requirements analysis can significantly improve process effectiveness [11], [6], including in terms of accuracy, time efficiency, and result consistency. Transformer-based models such as BERT and LLMs have been reported to outperform conventional methods in understanding semantic context [12], [13], [14]. Furthermore, hybrid approaches that combine multiple AI techniques have demonstrated superior performance across various requirements engineering tasks [15], [16]. Despite these advantages, the implementation of AI in requirements analysis still faces several challenges. These include data imbalance, limited model generalization across different domains and languages, and a lack of transparency in AI decision-making processes [17], [18], [19]. Moreover, most existing studies rely on academic datasets, resulting in limited validation

in real-world industrial environments [20], [21]. Based on these challenges, a comprehensive study is required to understand how AI integration can support automation and improve the effectiveness of requirements analysis in information systems. Therefore, this study employs a Systematic Literature Review (SLR) approach to identify the AI techniques used, evaluate their impact on the effectiveness of the requirements analysis process, and examine the challenges encountered in their implementation.

2. RESEARCH METHODS

This study employs a Systematic Literature Review (SLR) method [22] to systematically examine the use of Artificial Intelligence (AI) in information system requirements analysis. The SLR method was selected because it provides a comprehensive synthesis of relevant studies while minimizing bias in the study selection process. The research process follows the PRISMA 2020 guidelines [23] (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), which consist of the identification, screening, eligibility assessment, and inclusion stages.

2.1. Literature Search Strategy

The literature search was conducted across several indexed academic databases, including IEEE Xplore, ScienceDirect, SpringerLink, ACM Digital Library, and Scopus. The search was performed using a combination of keywords related to the research topic, such as “Artificial Intelligence”, “Machine Learning”, “Requirements Engineering”, “Requirements Analysis”, and “Automation”.

To obtain relevant results, the following Boolean search query was applied:

("Artificial Intelligence" OR "Machine Learning" OR "Deep Learning" OR "Large Language Model") AND ("Requirements Engineering" OR "Requirements Analysis") AND ("Automation" OR "Classification" OR "Extraction").

2.2. Inclusion and Exclusion Criteria

The inclusion criteria were defined to identify articles relevant to this study, as follows:

- a. The article is an empirical study that discusses the application of Artificial Intelligence (AI) in requirements analysis.
- b. The article was published between 2020 and 2025.
- c. The article is available in full-text and written in English.
- d. The article includes evaluation results using clear metrics, such as precision, recall, or F1-score.

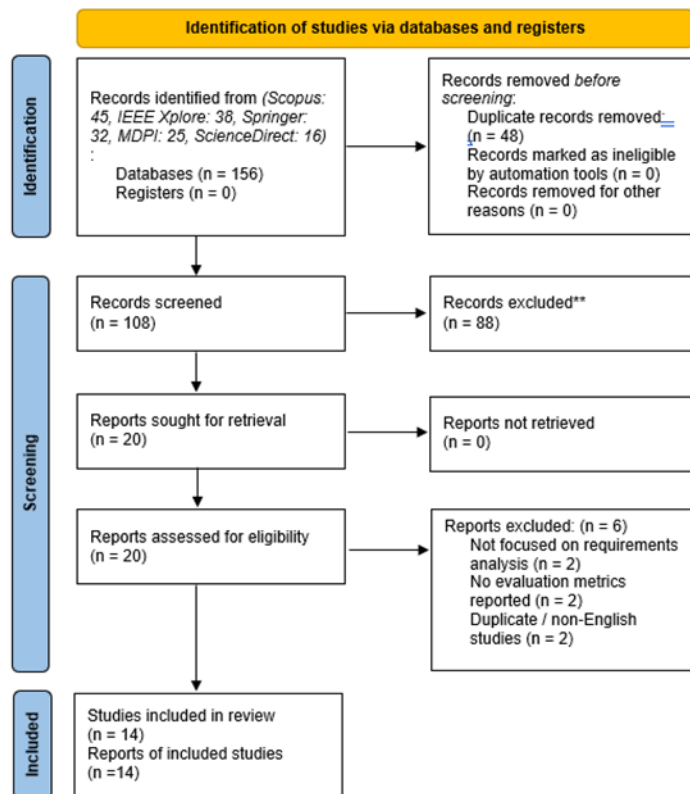
The exclusion criteria include:

- a. Articles that are opinion-based, non-systematic reviews, or do not include empirical validation.
- b. Articles that do not focus on information system requirements analysis.
- c. Duplicate articles or earlier versions of the same publication.

2.3. Study Selection Process

The study selection process was conducted systematically by following the PRISMA 2020 framework, which includes the identification, screening, eligibility, and inclusion stages. During the identification stage, articles were collected from various academic databases using predefined keywords, followed by the removal of duplicate records. The screening stage was performed by reviewing the titles and abstracts to assess their initial relevance to the research topic. Articles that met the initial criteria were then evaluated in the eligibility stage through full-text review to ensure their compliance with the predefined inclusion and exclusion criteria. Based on this process, a total of 14 articles were selected and included in the analysis. The complete study selection process is illustrated in Figure 1.

PRISMA 2020 flow diagram for new systematic reviews which included searches of databases, registers and other sources



*Consider, if feasible to do so, reporting the number of records identified from each database or register searched (rather than the total number across all databases/registers)

**If automation tools were used, indicate how many records were excluded by a human and how many were excluded by automation tools.

Source: Page MJ, et al. *BMJ* 2021;372:n71. doi: 10.1136/bmj.n71.

This work is licensed under CC BY 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>

Figure 1. PRISMA Flow Diagram of the Study Selection Process

2.4. Data Extraction

Data from each selected article were extracted using a structured form. The collected information includes the authors and publication year, research objectives, AI techniques used, applied methods, datasets utilized, evaluation metrics, main findings, and study limitations.

2.5. Data Synthesis

The extracted data were analyzed using a thematic synthesis approach. The analysis was conducted by grouping the studies based on the types of AI techniques employed and the tasks in requirements analysis, such as classification, ambiguity detection, requirements extraction, and prioritization. Furthermore, the effectiveness of AI implementation was evaluated based on the metrics reported in each study, including accuracy, precision, recall, and F1-score. The results of this synthesis were used to address the research questions and to identify the strengths and limitations of each approach.

3. RESULTS AND DISCUSSION

3.1. Characteristics of the Analyzed Studies

Based on the selection results using the Systematic Literature Review (SLR) method, a total of 14 articles met the inclusion criteria and were included in this study. These articles were published between 2020 and 2025 and were sourced from various reputable international journals.

Most studies focus on the application of Artificial Intelligence (AI) in various requirements engineering sub-tasks, such as requirements classification (functional and non-functional requirements), ambiguity detection, requirements extraction from unstructured documents, traceability link recovery, and requirements prioritization [6], [7]. In terms of datasets, the majority of studies utilize text-based data derived from software requirement documents, such as Software Requirement Specifications (SRS), user stories, and other project documents. The datasets vary from publicly available datasets, such as PROMISE and Dronology, to real-world industrial datasets, although their usage remains limited [11], [18].

3.2. Artificial Intelligence Techniques Used

The analysis results indicate that various Artificial Intelligence (AI) techniques have been applied in requirements analysis, which can be categorized into several main groups.

- a. **Machine Learning (ML)**
Machine Learning techniques such as Support Vector Machine (SVM), Random Forest, Decision Tree, and K-Nearest Neighbors (KNN) have been widely used in earlier studies for requirements classification tasks. This approach generally relies on feature extraction techniques such as TF-IDF and requires manual feature engineering processes [4], [15].
- b. **Deep Learning (DL)**
Subsequent developments show the use of Deep Learning techniques, such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM), which are capable of capturing sequential patterns in textual data without requiring extensive manual feature engineering [11].
- c. **Transformer dan Large Language Models (LLM)**
Transformer-based models such as BERT, RoBERTa, and various Large Language Models (LLMs) have become the dominant approach in recent studies. These models are capable of understanding semantic context more effectively, resulting in improved performance across various requirements analysis tasks [6], [7].
- d. **Hybrid and Ensemble Approaches**
Several studies have developed hybrid approaches by combining multiple AI techniques, such as integrating LLMs with sequential models (LSTM/GRU) or combining machine learning with logical reasoning. These approaches have been shown to improve performance compared to single-method approaches [15], [16].

3.3. Impact of AI on Requirements Analysis Effectiveness

- a. **Accuracy Improvement**
Most studies report a significant improvement in accuracy in requirements analysis. Transformer-based models demonstrate the best performance, with F1-scores ranging from 0.85 to 0.95 [6], [11]. In addition, the use of techniques such as data fusion and ensemble learning has been shown to further enhance model performance compared to conventional approaches [15].
- b. **Time Efficiency**
The application of Artificial Intelligence (AI) enables the automation of requirements analysis processes, thereby reducing the time required compared to manual methods. Tasks that previously required substantial time, such as requirements classification and extraction, can be completed more efficiently using AI models [9].
- c. **Result Consistency**
AI also improves the consistency of analysis results by reducing dependence on human subjectivity. Trained models can produce more stable and consistent decisions compared to human analysts, who may have varying interpretations [4].

3.4. Challenges in the Application of Artificial Intelligence

Despite its advantages, the application of Artificial Intelligence (AI) in requirements analysis still faces several challenges that need to be addressed.

First, data imbalance (class imbalance) is a major issue in requirements classification, particularly for non-functional requirements, which typically have fewer data instances compared to functional requirements [17].

Second, limited model generalization is also a significant challenge, especially when models are applied to datasets with different domains or languages from the training data [6]. This may lead to a decline in model performance when used in different contexts.

Third, the lack of transparency or explainability in AI models, particularly in deep learning and Large Language Models (LLMs), makes it difficult for users to understand the reasoning behind the generated decisions [18].

In addition, most studies still rely on academic datasets, resulting in limited validation in real-world industrial environments. This indicates the need for further research to evaluate the reliability of AI models in practical, real-world scenarios [21].

3.5. Synthesis and Implications

Based on the analysis of the reviewed studies, it can be concluded that the application of Artificial Intelligence (AI) in information system requirements analysis has experienced significant development. Approaches that were initially dominated by conventional machine learning techniques have evolved toward the use of deep learning models, particularly Transformer-based models and Large Language Models (LLMs), which are capable of understanding semantic context more effectively.

In addition, hybrid approaches that combine multiple AI techniques show an increasing trend in performance across various requirements engineering tasks. These combined methods are able to overcome the limitations of individual

approaches while improving the accuracy and efficiency of the requirements analysis process. The implications of this study indicate that AI integration has significant potential to be adopted in requirements analysis processes within industrial environments. However, its implementation should consider several important aspects, such as data quality and availability, model generalization capability, and interpretability, to ensure that the results are reliable and can be effectively utilized. Therefore, further development in AI integration, particularly in adaptive and context-aware approaches, is essential to support broader implementation in software engineering practices.

3.6. Discussion

The findings of this Systematic Literature Review (SLR) highlight the significant evolution of Artificial Intelligence (AI) techniques in supporting requirements analysis within information systems. The transition from traditional machine learning methods to deep learning and Transformer-based models demonstrates a clear shift toward approaches that are more capable of capturing semantic complexity in textual data. Transformer-based models and Large Language Models (LLMs), in particular, consistently achieve higher performance, with F1-scores reaching up to 0.95, indicating their superiority in tasks such as classification, extraction, and ambiguity detection.

However, despite these advancements, the results also reveal several critical limitations. One of the main challenges is the issue of data imbalance, especially in classifying non-functional requirements, which often leads to biased model performance. Additionally, the lack of generalization across different domains and languages suggests that many models are still highly dependent on the characteristics of the training dataset. This limitation reduces their applicability in real-world industrial environments, where data diversity is significantly higher. Another important concern is the lack of explainability in advanced AI models, particularly deep learning and LLM-based approaches. While these models offer high accuracy, their “black-box” nature limits transparency, making it difficult for analysts to trust and validate the results. This is especially critical in requirements engineering, where decision traceability is essential.

Furthermore, although hybrid and ensemble approaches demonstrate improved performance, they also introduce additional complexity in model design and implementation. This trade-off between performance and complexity must be carefully considered, particularly in practical applications where computational resources and interpretability are important factors. Overall, this study indicates that AI has strong potential to enhance the effectiveness of requirements analysis by improving accuracy, efficiency, and consistency. However, future research should focus on developing more interpretable, generalizable, and domain-adaptive models, as well as validating these approaches using real-world industrial datasets to ensure their practical applicability.

4. CONCLUSION

This study conducted a Systematic Literature Review (SLR) to examine the application of Artificial Intelligence (AI) in information system requirements analysis. The findings indicate that AI has significantly improved the effectiveness of the requirements analysis process, particularly in terms of accuracy, time efficiency, and result consistency. Advanced approaches, especially Transformer-based models and Large Language Models (LLMs), demonstrate superior performance in handling complex semantic tasks such as requirements classification, ambiguity detection, and information extraction. However, despite these advantages, several challenges remain. Issues such as data imbalance, limited model generalization across different domains and languages, and the lack of explainability in AI models continue to hinder broader adoption. In addition, most studies still rely on academic datasets, which limits the validation of AI models in real-world industrial environments. These limitations highlight the need for further research to improve the robustness and applicability of AI-based approaches.

Overall, this study confirms that AI integration has strong potential to transform requirements analysis processes into more automated, efficient, and consistent practices. Future research should focus on developing more interpretable and adaptive models, incorporating diverse and real-world datasets, and exploring hybrid approaches that combine multiple AI techniques. Such advancements are essential to ensure that AI-based solutions can be effectively implemented in practical software engineering environments.

REFERENCES

- [1] C. Arora, J. Grundy, and M. Abdelrazek, “Advancing requirements engineering through generative AI: Assessing the role of large language models,” in *Lecture Notes in Computer Science*, 2024, pp. 129–148.
- [2] I. Albana, F. Fadhillah, L. Widiastuti, and M. Veda Yesenia, “Analisis Faktor Kesuksesan Proyek Perangkat Lunak,” *J. Informatics Interact. Technol.*, vol. 1, no. 3, pp. 148–156, 2024.
- [3] S. Winkler and J. von Pilgrim, “A survey of traceability in requirements engineering and model-driven development,” *Softw. Syst. Model.*, vol. 9, no. 4, pp. 529–565, 2010.
- [4] T. Li, S. Wang, D. Lillis, and Z. Yang, “Combining machine learning and logical reasoning to improve requirements traceability recovery,” *Appl. Sci.*, vol. 10, no. 20, 2020.
- [5] N. Condori-Fernandez, M. Daneva, R. Wieringa, and O. Dieste, “Stakeholder-driven requirements analysis: Challenges and future directions,” *Requir. Eng.*, vol. 25, no. 4, pp. 383–401, 2020.

- [6] F. Yucalar, “Developing an advanced software requirements classification model using BERT: An empirical evaluation study,” *Appl. Sci.*, vol. 13, no. 20, p. 11127, 2023.
- [7] S. Saleem, M. Asim, and A. Dengel, “ReqNet: An LLM-driven computational framework for automated requirements extraction from unstructured documents,” *Complex Intell. Syst.*, vol. 12, no. 1, 2026.
- [8] L. Zhao et al., “Natural language processing for requirements engineering: A systematic mapping study,” *arXiv Prepr. arXiv2004.01099*, 2020.
- [9] A. Radwan, M. Abdel-Fattah, and W. Mohamed, “Smart agile prioritization and clustering: An AI-driven approach for requirements prioritization,” *IEEE Access*, vol. 13, pp. 127335–127350, 2025.
- [10] A. Ferrari, G. Spagnolo, and S. Gnesi, “AI-based techniques for requirements prioritization: A review,” *IET Softw.*, vol. 15, no. 1, pp. 1–29, 2021.
- [11] K. Rahman, A. Ghani, S. Misra, and A. Rahman, “A deep learning framework for non-functional requirement classification,” *Sci. Rep.*, vol. 14, no. 1, 2024.
- [12] J. Devlin, M. Chang, K. Lee, and K. Toutanova, “BERT: Pre-training of deep bidirectional transformers for language understanding,” in *NAACL-HLT 2019*, 2019, pp. 4171–4186.
- [13] T. Brown, B. Mann, N. Ryder, M. Subbiah, and J. Kaplan, “Language models are few-shot learners,” *Adv. Neural Inf. Process. Syst.*, 2020.
- [14] W. Zhang, Z. Yang, Y. Li, and S. Xu, “Transformer-based models for requirement classification: A comparative study,” *IEEE Access*, vol. 8, pp. 211225–211239, 2020.
- [15] M. Raza, V. Mir, J. Rasheed, M. Yesiltepe, and S. Alsubai, “Enhancing software requirement classification via dataset fusion and machine learning,” *Acta Infologica*, 2025.
- [16] M. Sharma, S. Sabharwal, and R. Sibal, “Hybrid AI models in software requirements engineering,” *Softw. Qual. J.*, vol. 29, no. 2, pp. 269–288, 2021.
- [17] P. Limaylla-Lunarejo and J. A. Pow-Sang, “Improving the multi-class classification of non-functional requirements in Spanish,” *IEEE Lat. Am. Trans.*, vol. 18, no. 4, pp. 743–751, 2020.
- [18] S. Bashir et al., “Requirements ambiguity detection and explanation with large language models: An industrial study,” in *2025 IEEE International Conference on Software Maintenance and Evolution (ICSME 2025)*, 2025, pp. 620–631.
- [19] V. Gupta, J. M. Fernandez-Crehuet, C. Puente, and R. Gupta, “Explainable AI in requirements engineering: Challenges and opportunities,” *IEEE Softw.*, vol. 38, no. 3, pp. 86–92, 2021.
- [20] T. Hey, D. Fuchs, J. Keim, and A. Koziolk, “Requirements traceability link recovery via retrieval-augmented generation,” in *Lecture Notes in Computer Science (LNCS)*, vol. 13588, 2025, pp. 381–391.
- [21] D. Romano, K. Feichtinger, D. Beuche, U. Ryssel, and R. Rabiser, “Bridging the gap between academia and industry in AI-driven requirements engineering,” in *2025 IEEE International Conference on Software Maintenance and Evolution (ICSME 2025)*, 2025, pp. 620–631.
- [22] B. Kitchenham and S. Charters, “Guidelines for performing systematic literature reviews in software engineering,” *EBSE Tech. Rep.*, pp. 1–57, 2007.
- [23] M. Page et al., “The PRISMA 2020 statement: An updated guideline for reporting systematic reviews,” *Med. Flum.*, vol. 57, no. 4, pp. 444–465, 2021.