24th International Conference on Construction Applications of Virtual Reality (CONVR 2024)

4-6 November 2024, Western Sydney University, NSW, Australia

Risk Assessment for Construction Safety with Retrieval-augmented Generation and Agentic Artificial Intelligence

Kilian Speiser[1*], Guido Maciocci[2], and Jochen Teizer[1]

[1]Technical University of Denmark (DTU), Kongens Lyngby, Denmark
[2]AecFoundry, Rome, Italy

* kilsp@dtu.dk

Abstract. Construction sites represent hazardous work environments, necessitating comprehensive risk assessments to ensure worker safety. Traditional methods of risk assessment are critical for identifying and mitigating potential hazards. However, the complex nature of these safety documents often poses a significant challenge. This gap in safety knowledge can lead to inadequate risk assessments and heightened workplace hazards. The emergence of advanced technologies in Artificial Intelligence (AI) such as Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG) presents new opportunities to address these challenges. While LLMs enhance natural language processing, this hallucination and lack of reliable source referencing can undermine their effectiveness. RAG offers a promising solution by integrating information retrieval, though its application in construction safety engineering remains unexplored. This paper introduces a safety risk assessment platform leveraging agentic AI systems, LLMs, hybrid semantic search, and RAG to generate expert-level safety risk assessments. The system utilized extensive safety knowledge sources from Safe Work Australia into a framework where users can create risk assessment reports for specific work tasks purely based on natural language input. Preliminary evaluations of the platform demonstrate its effectiveness in generating such reports, although improvements are necessary to reduce document verbosity and enhance usability.

Keywords: Retrieval-Augmented Generation (RAG), Agentic Artificial Intelligence, Construction Safety, Risk Assessment, Safety Management, Natural Language Processing (NLP).

1 Introduction

Construction sites offer hazardous work environments and require a complex risk assessment followed by a detailed safety concept before tasks can start [1]. Effective safety risk assessments, known as Sicherheit- und Gesundheitsschutzkoordination in Germany (SiGeKo), Safe Work Method Statements (SWMS) in Australia, Risk Assessment Method Statements (RAMS) in the UK, and Job Safety Analyses (JSA) in the US, are crucial for identifying potential hazards early in the construction process. These assessments require organizations, including construction companies, to analyze their on-site work in detail, document potential risks, and implement control measures to mitigate these risks before the project starts and before any work is carried out.

This version of the contribution has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements or corrections. The Version of Record will be published in the Proceedings of the CONVR 2024 conference, and a DOI link will be provided once available.

These underlying concepts are similar in each country and refer to a variety of knowledge inputs provided by laws, regulations, and best practices. Available documents that guide this process typically comprise a tremendous amount of textual information, making it often a time-consuming, manual and error-prone task to aggregate such information in a project-specific context. However, many subcontractors in the construction industry are small to medium-sized businesses and may have limited safety resources and expertise. This gap in resources and safety knowledge can lead to inadequate risk assessments and increased workplace hazards. As the accident records in the construction industry remain high [2,3], these tasks can be one of the contributing factors.

Further complicating matters is the labor shortage in the construction industry, which is causing a demand for safety professionals [4]. These experts are vital for interpreting and implementing regulations effectively. However, even when available, these experts may struggle to keep up with evolving safety standards, resulting in knowledge gaps and oversight issues. Additionally, less experienced safety professionals may lack the necessary skills to navigate these complexities independently, placing additional strain on safety resources.

With the emergence of Large Language Models (LLM), performing a variety of Natural Language Processing (NLP) tasks, such as browsing and assembling knowledge, has become more efficient [5]. However, LLMs have two undesired characteristics. First, they risk hallucinating, a phenomenon where the LLM invents answers that are wrong, and second, LLMs do not provide reliable references to the source of the cited knowledge [6]. LLMs have been adopted for construction safety purposes [7,8]. Retrieval-augmented generation (RAG) as a specification of Generative AI is a promising solution to these issues [9], but it has not been studied in the case of construction, particularly not for construction safety engineering and safety risk assessments for construction sites.

While RAG and LLM can generate natural language from diverse input, another challenge in creating structured feedback in the form of a SWMS is that technical hurdles exist, but agentic AI provides a method to overcome those. No existing commercial solution exploits these capabilities.

This paper proposes the safety risk assessment platform SWMS AI to address this critical issue by leveraging advanced agentic AI systems and a curated knowledge base on workplace safety. The platform provides access to safety expertise, enabling users with little knowledge of workplace safety practices to create expert-level safety risk assessments for any high-risk activity. By making safety expertise universally accessible, SWMS AI empowers small to medium-sized businesses to enhance their safety practices, ultimately contributing to a safer construction industry. Integrating advanced technologies like LLMs and agentic AI systems can significantly enhance safety protocols, offering real-time, data-driven risk assessments and safety management solutions that are both efficient and comprehensive.

2 Research Methods and Objectives

This research employs a qualitative approach to develop a RAG-enabled tool for safety professionals to create risk assessment statements. It began with a comparison of existing tools for safety management in construction sites, iteratively clustering them by various features they support. The focus was on identifying their use of AI, particularly generative AI. This comparison revealed a missing component in these tools: supportive features to expedite risk assessments by providing relevant knowledge.

Based on this identified need, the next step involved developing the SWMS AI tool, designed to support safety professionals in creating task-specific risk assessment documentation. Given the study's focus on the Australian construction industry, the requirements adhere to Australian SWMS standards.

Creating these statements requires significant effort and diverse knowledge. The proposed RAG-enabled AI solution SWMS AI browses large knowledge bases and provides informative responses in natural language to support safety experts in creating such statements. The development of the tool utilized several libraries, including the RAG library LangChain. In addition to providing natural language responses, another component, the SWMS generator, was developed to structure these responses into predefined groups: work sequences, hazards, and personal protective equipment (PPE). This structured response is facilitated by accessing the RAG-enabled system through three agents (Agentic AI).

The system was evaluated qualitatively by showcasing the risk assessment results for three tasks chosen across trades to allow for a general discussion not limited to singular tasks. The study qualitatively discusses the implications and benefits of such a system. Although in a preliminary stage, future work must quantitatively assess the system's reliability, accuracy, and usability in practice, requiring broader tests. Nevertheless, this study outlines several benefits and implications of RAG and Agentic AI for construction safety.

3 Relevant Existing Tools for Construction Safety Management

The study started with an evaluation of existing safety tools for construction safety management. The features of the 11 tools were compared and clustered as listed in Table 1.

Table 1. Existing safety management tools and feature comparison.

	Protex.ai	Intenseye	newmetrix	iAuditor	Enablon	Quentic	Hammertech	Intelex	ChatSafeAI	Procore	iSafe	Safe site
Commercial	x	x	x	x		x	x	x		x		x
Academic									x		x	
Reporting	x	x	x	x								x
Event log	x			x								x
Data-driven decision making	x	x	x		x	x	x	x				x
Real-time event capture	x	x										
Predictive safety analysis			x		x							
Checklists for inspections				x	x	x						x
Access control							x					
Training				x			x					
Quality management								x				
ESG reporting and management					x	x		x				
Mobile solution				x	x	x	x	x				x
Camera integration	x	x										

Among all tools, five – Protex.ai, Intenseye, Newmetrix, Safe site, and iAuditor – provide essential reporting features for compliance and tracking. Protex.ai and iAuditor also include event logging, which is useful for maintaining detailed incident records. Data-driven decision-making is supported by seven tools, emphasizing the importance of analytics in modern safety management. Only Protex.ai and Intenseye support real-time event capture, which is crucial for immediate response and thus integrates with cameras.

Predictive safety analysis, which helps anticipate potential issues, is offered by iAuditor and Enablon, although this feature is rare. Hammertech stands out for managing access control, suggesting a specialized feature. Intelex and Procore offer essential training for workforce preparedness. Enablon is noted for quality management, focusing on maintaining standards. Enablon, Quentic, and Hammertech support ESG reporting, reflecting a growing focus on sustainability. Most tools include mobile solutions, which are crucial for accessibility and field operations on sites.

The comparison of existing solutions highlights the rarity of AI integration. Protex.ai and Intenseye use computer vision with integrated camera systems to capture relevant events in real-time. Newmetrix and Enablon mention predictive features using AI capabilities. However, none of the investigated tools

This version of the contribution has been accepted for publication, after peer review but is not the Version of Record and does not reflect post-acceptance improvements or corrections. The Version of Record will be published in the Proceedings of the CONVR 2024 conference, and a DOI link will be provided once available.

utilize generative AI or RAG. This finding has prompted the authors to propose a RAG-enabled tool for risk assessment, where an extensive knowledge base is queried efficiently, and relevant context is returned in a structured form as SWMS for safety experts, as described in the following Section.

4 Safety Risk Assessment from RAG and Agentic AI

4.1 Overview

Since the following text refers to Australia, we call our safety risk assessment tool SWMS AI from hereon. The developed SWMS AI platform [10] employs a sophisticated technological framework to deliver expert-level safety risk assessments. This framework, illustrated in Fig. 1, comprises four key components: LLMs, Hybrid Semantic Search, AI Agents, and RAG. The relevant safety knowledge is stored in a vector database by transforming the textual information into vectors using an embeddings model. Once the user interacts with the system through the user interface, the query is reformulated into three separate queries, and hybrid semantic search provides the most relevant context from the vector database. The user's query and the relevant documents are forwarded to three agents, each of them using this input and possibly adding output from other agents to provide context for the SWMS by utilizing their LLM capabilities. The generative responses are returned to the user as SWMS. The concept of first retrieving the relevant context before requesting the LLM to respond follows the RAG principle.

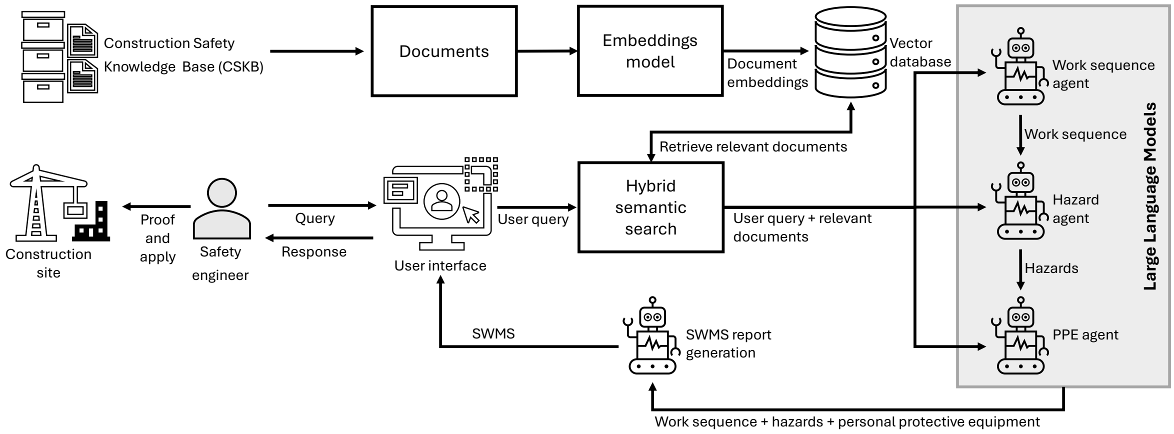

Fig. 1. The developed system for generating Safe Work Method Statements using LLM, Hybrid Semantic Search, AI Agents, and RAG.

By leveraging these components, SWMS AI provides an accessible and efficient tool for generating high-quality safety risk assessments, empowering users with limited safety expertise to implement robust safety practices in high-risk industries. The following sections describe the individual components and processes of SWMS AI in detail.

4.2 Building the Construction Safety Knowledge Base

The foundation of the SWMS AI platform's effectiveness lies in its extensive Construction Safety Knowledge Base (CSKB), constructed from publicly available safety information. As this research is targeted at Australian risk assessment, the knowledge was sourced from Safe Work Australia. This knowledge base includes over a thousand documents covering a wide range of workplace safety topics. These documents are authoritative and provide a robust foundation for the risk assessment capabilities.

Converting the knowledge base into a vectorized format is essential for seamless integration with the RAG architecture. The knowledge was first collected and then converted into vectors. To collect the knowledge in textual format, an algorithm developed using Apify crawled the web pages of Safe Work Australia and extracted the text from all subpages into markdown files. Parallel, PDF documents were downloaded from Safe Work Australia and parsed with LlamaParse to output structured markdown text.

In the second step, the textual information in each document was separated into chunks. This text splitting is based on markdown syntax (e.g., subheadings and paragraphs stay together) with a max chunk size of 1000 tokens and an overlap of 200 tokens. These chunks were then converted into vectors using OpenAI's embedding model *text-embedding-ada-003*. The model converts the textual data into a high-dimensional space where semantic similarities are represented by proximity. The conversion resulted in a total of approximately 20,000 vectors. These vectors are finally indexed in the vector database Azure Cognitive Search Index.

Ensuring comprehensive coverage and maintaining the integrity of context is a significant aspect of this process. By creating overlapping chunks, the system mitigates the risk of losing critical context that might span across document boundaries. Additionally, the use of advanced embedding models and vector databases allows the system to perform optimized semantic and vector searches, thereby enhancing the quality and reliability of the generated safety assessments.

4.3 Knowledge Retrieval and Response Generation

The retrieval of the relevant knowledge and the generation of the response is divided into two steps. First, the relevant context is retrieved using a multi-query retriever and hybrid-semantic search. This approach combines traditional keyword search with advanced vector-based semantic search. It ensures precise and contextually relevant information retrieval from the platform's extensive knowledge base [11,12], enhancing the accuracy of hazard identification and control measure suggestions. Second, the relevant documents and the user's query are forwarded to several agents to generate the SWMS. These specialized AI agents are engineered through prompt templates and handle distinct tasks within the risk assessment process. These agents break down complex queries into manageable tasks, leveraging the knowledge base for informed outputs [13].

The hybrid semantic search approach illustrated in Fig. 2 first reformulates the user's query into three variations using the LLM GPT-4o. This concept is known as multi-query-retrieval and enhances the precision of the search process by breaking down the complex input into more targeted queries [14]. Parallel, the keyword search also retrieves the most relevant documents. Combining the results from both approaches, a reciprocal rank function and a semantic re-ranker determine the final selection of documents representing the most relevant context based on the initial user query. These top n results are eventually forwarded to three different agents.

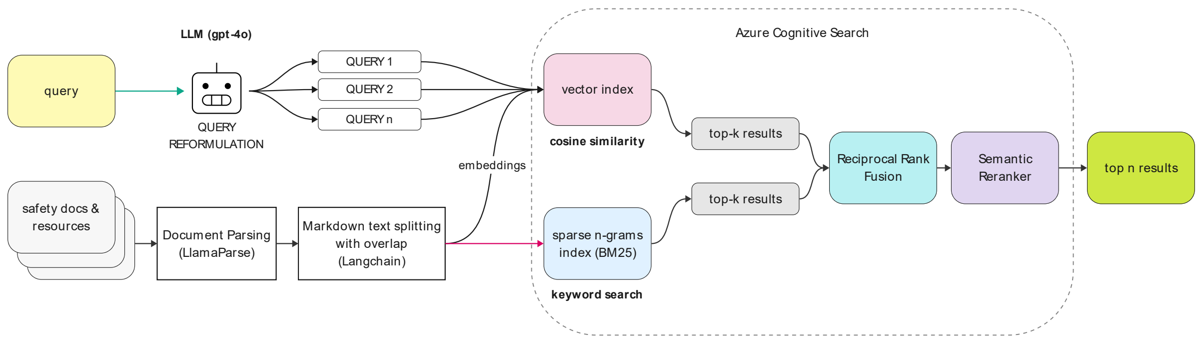

Fig. 2. Retrieving the most relevant documents from the knowledge base using hybrid semantic search.

SWMS AI includes three agents, each of them using different inputs and prompt templates to generate the relevant output for their purpose (compare Fig. 1). The *Work Sequence Agent* constructs a logical sequence of work steps, including pre-and post-work activities, based on the user's input and the relevant context provided by the hybrid semantic search. It accesses the knowledge base to ensure the sequence is informed by relevant safety protocols and standards, resulting in a detailed description of each specific activity. The *Hazard Agent* identifies potential hazards for each work sequence step. It retrieves relevant knowledge from the database to list hazards, assign risk scores, and suggest control measures. Using a risk matrix of severity versus frequency, the agent applies a hierarchy of controls to determine the most appropriate mitigation strategies. To generate an adequate response, the hazard agent also uses the output from the work sequence agent. In the last step, the *PPE Agent* identifies the required specialized equipment and PPE. This agent analyzes the generated work sequence and the original user description. It ensures all necessary equipment is accounted for. Hence, it utilizes the user's query, the most relevant documents from the search, and the output from the hazard agent to generate this content. The final response of each agent is then aggregated and forwarded to the SWMS generator component.

4.4 SWMS Report Generation

By iteratively analyzing and processing each step of the work sequence, the system constructs a comprehensive safety risk assessment. This approach, while increasing response latency, significantly improves the accuracy, precision, and relevance of the generated assessments. The structured results provided by the agentic AI framework are eventually integrated into a detailed SWMS as a PDF document. While this study targets the Australian market, a similar approach should also be applicable to other regional or national requirements.

The created documents include (a) the description of the job, (b) the required equipment for performing the job, (c) the required personal protective equipment, and (d) a list of all hazards for each step in the work sequence. Each sequence in the task is listed with the identified hazards, a risk score, and the suggested control measures. The following Section overviews three excerpts from the generated SWMS and qualitatively discusses the results.

5 Evaluation

This Section shows the results of the tool in three examples. First, a conversation with the system when asking for heat-related hazards is depicted. Second, two excerpts from the generated SWSM are shown and discussed.

5.1 Example A: Heat Hazards

A core feature of the system involves generating natural language responses based on user-posed questions. Initially, we will illustrate the system's responses, followed by examples that analyze these responses for a comprehensive risk assessment report.

In the first example, the system addresses hazards associated with hot work on-site (Fig. 3). SWMS AI identifies relevant hazards, including fire, explosion, burns, and exposure to extreme heat. When prompted about mitigation measures, the system organizes them according to the hierarchy of controls [15]. This concept prioritizes eliminating and substituting hazards, followed by implementing engineering and administrative controls before suggesting personal protective equipment (PPE). The SWMS AI structures the controls according to these categories, as the prompt was designed for this purpose.

Looking at the engineering controls in Fig. 3, the response includes all relevant options: Cooling devices, physical barriers, and automated equipment. The response furthermore suggests adequate administrative controls: training, supervision, and the planning of tasks during lower-temperature times. This example indicates that the system can generate accurate answers based on the provided input, but the next examples will underline that the SWSM AI tool can also structure the responses and generate meaningful risk assessments.

User

What are the hazards associated with hot works on site?

SWMS AI

The hazards associated with hot works on a site can include risks such as **fire, explosion, burns, and exposure to extreme heat**. These hazards can arise from various sources such as **welding, cutting, brazing, and the use of hot work equipment** It's essential to implement control measures to protect workers from these hazards, such as isolating workers hot items, marking or labeling hot equipment, and appropriate personal protective equipment (PPE) to avoid contact with heated surfaces. Understanding and mitigating these hazards is crucial for ensuring the safety and well-being of workers performing hot works on a site.

User

What are some control measures that we can implement?

SWMS AI

1. Try to eliminate the need for hot works by using alternative methods. For example, use mechanical fastening instead of welding.
2. Substitute the Hazard: Replace hazardous hot work processes with safer alternatives. For instance, substitute hold cutting techniques with hot cutting techniques
3. Engineering Controls
   Ventilation: Ensure good ventilation to remove heated air or steam from hot processes.**Cooling Devices:** air conditioning or fans improve air flow and reduce heat.**Physical Barriers:** Shields, barriers, and guards isolate hot machinery or surfaces from workers.**Automated Equipment:** Use automated or remote-controlled machinery to perform hot works, reducing the need for direct human involvement.
4. Administrative Controls
   Work Scheduling: Schedule hot works during cooler parts of the day.**Training:** Provide comprehensive training **Supervision:** Ensure proper supervision during hot works.**Permits:** Implement a hot work permit system
5. Personal Protective Equipment (PPE)
   Protective Clothing: flame-resistant clothing, gloves, and face shields.**Hydration**: Provide cool drinking water and encourage regular hydration. **Heat Stress Monitoring**: Monitor workers for signs of heat stress
6. Emergency Preparedness
   Fire Safety: Have firefighting equipment readily available.**First Aid:** Provide first aid training and ensure first aid kits are accessible

Fig. 3. The conversation with the system provides specific heat-related hazards based on the user's input.

5.2 Example B: Positioning a Loader for Materials Loading

The second example addresses the positioning of a truck loader for loading materials from a construction site. Fig. 4 shows an excerpt of the generated SWMS, which includes three typical hazards: collisions, falling objects, and equipment failure. The document also suggests relevant PPE: head protection, safe footwear, and protective clothing. The example demonstrates that the information is structured in a tabular format, allowing stakeholders to identify relevant areas easily. Additionally, the risk score helps mitigate risks effectively by providing a metric for prioritization.

Position the Loader

Position the loader near the material pile, ensuring a safe distance from any obstacles or workers.

Hazard	Risk Score	Controls
Collision with obstacles or workers	15	- Ensure a safe distance is maintained from obstacles and workers - Use spotters to guide the loader operator and ensure safe positioning - Implement traffic control measures to prevent unauthorized access to the work area - Provide training to loader operators on safe positioning techniques
Falling objects	12	- Inspect the material pile for stability and remove any loose or unstable material - Use proper lifting techniques to minimize the risk of material falling during positioning - Ensure the loader's forks is in good working condition and securely grips the material - Wear appropriate personal protective equipment (PPE), e.g., safety vest and hard hat
Equipment failure	10	- Perform regular maintenance and inspections of the loader to identify and address any potential issues, e.g., ensure all safety features are in working condition - Train loader operators on equipment operation and maintenance procedures - Have a backup loader available in case of equipment failure

Personal Protective Equipment

Fig. 4. Excerpt of the SWMS for positioning a loader as one step in the task of loading materials.

5.3 Example C: Chainsaw Use

The last example involves using a chainsaw. The complete SWMS spans several pages, so this Section focuses on one step in the work sequence: starting the chainsaw.

Table **2** lists all relevant hazards for this task and the suggested PPE. For clarity, a table was recreated from the generated PDF report.

Table 2. Risk assessment generated for the step of starting a chainsaw provided by SAWS AI.

Hazard	Risk Score	Controls	Personal Protective Equipment
Severe lacerations	25	• Wear cut-resistant gloves • Ensure the chainsaw is properly maintained and the saw chain is sharpened to the manufacturer's specifications • Maintain a safe working distance from other people • Ensure the saw chain is clear of obstructions	 Head Protection  Safety Footwear  Eye Protection  Protective Gloves  Protective Clothing  High-vis Clothing
Chainsaw kickback	25	• Ensure chainsaw is properly maintained and the saw chain is tensioned correctly • Use a chainsaw with a shorter guide bar length for better control • Maintain a safe working distance from other people • Ensure the saw chain is clear of obstructions	
Eye injuries from wood fragments	20	• Wear safety goggles or a face shield • Maintain a safe working distance from other people • Ensure the saw chain is clear of obstructions	
Being struck by falling trunks, branches, and other objects	20	• Ensure the area is clear of people and objects that could be struck • Establish a drop zone for falling trunks and branches • Use appropriate felling techniques to control direction of fall • Maintain a safe working distance from other people	
Musculoskeletal disorders	15	• Use proper lifting and carrying techniques • Take regular breaks to avoid overexertion • Use ergonomic equipment and tools • Maintain a safe working distance from other people	
Exposure to high levels of noise and hand/arm vibration	15	• Wear hearing protection • Take regular breaks to reduce exposure to vibration • Use anti-vibration gloves • Maintain a safe working distance from other people	
Fire from fuel source of chainsaw	10	• Ensure the chainsaw is fueled/refueled in a safe designated area • Store fuel in approved containers • Avoid smoking or open flames near the fuel source • Maintain a safe working distance from other people	
Cutting life support systems whilst working aloft and falling	20	• Ensure proper training and certification for working at heights • Use appropriate fall protection equipment • Secure the chainsaw to the operator/platform with a tool lanyard • Maintain a safe working distance from other people	
Burns from hot exhausts	10	• Avoid touching hot exhausts • Allow the chainsaw to cool down before handling • Maintain a safe working distance from other people	
Fumes when working in confined spaces	10	• Ensure proper ventilation in confined spaces • Use respiratory protection if necessary • Maintain a safe working distance from other people	

Several controls are repeated across multiple hazards, highlighting their general importance, e.g., maintaining a safe working distance and ensuring the chain is clear of obstructions. However, the repetition of the same controls extends the document and may overwhelm readers or diminish its usability. A condensed version could improve the SWMS's usability and increase worker acceptance.

The risk scores emphasize the most significant hazards. For example, wearing cut-resistant gloves is crucial due to their high probability of preventing injury. Conversely, a fire caused by the chainsaw's fuel, although severe, has a lower probability and, thus, a lower overall risk score.

6 Implications and Benefits

The construction industry remains one of the most hazardous sectors, with persistently high accident rates. One of the key issues identified is the complexity and volume of textual information that needs to be processed to develop effective safety risk assessments, such as SWMS in Australia, RAMS in the UK, and JSA in the US. Small to medium-sized enterprises particularly struggle due to their limited safety resources and expertise, often resulting in inadequate risk assessments and increased workplace hazards.

The comparison of existing safety management tools revealed that while many tools offer features like reporting, event logging, data-driven decision-making, and mobile solutions, they lack integration of advanced AI capabilities, particularly generative AI and RAG. Tools like Protex.ai, Intenseye, Newmetrix, and iAuditor incorporate elements of AI, such as computer vision and predictive analysis, but they do not fully utilize generative AI to streamline the creation of safety documents or enhance the contextual relevance of the information retrieved.

The proposed SWMS AI platform addresses these critical gaps by leveraging advanced technologies like LLMs, hybrid semantic search, AI agents, and RAG. This platform allows for access to expert-level safety risk assessments, making it possible for users with less safety expertise to generate comprehensive and accurate safety risk assessment reports.

The preliminary evaluation of the SWMS AI tool through qualitative analysis with safety experts and researchers indicates several potential benefits. First, the accuracy of the provided information is high, as the demonstrated examples indicate. Yet, the provided information may be too detailed as the example of the chainsaw starting illustrates where 10 hazards were identified and 37 corrective measures were proposed. Some of these controls are repetitive. For that reason, future improvements should consider a more concise documentation of controls. Nevertheless, an initial SWSM can be generated within minutes just by providing the context. Specialized prompt engineering training could enhance reliability and increase efficiency, but further research must validate that.

While the initial results are promising, future research directions should broader test the framework and quantitatively assess the system's reliability, accuracy, and usability in practice. Future studies should integrate the system with other tasks in construction safety, such as safety inspections on site. An integration into virtual construction safety training could enable trainees to receive real-time feedback in a virtual environment. Once such environments rely on real-world data, as proposed in previous research, the feedback could complete the assessment of the trainees in such systems [16,17]. Such a system could also be part of a digital twin for construction safety [18,19] for efficiently querying large amounts of knowledge and providing it in a structured format to the end user.

The results of this study indicate the exorbitant potential of technologies like RAG or Agentic AI for the construction industry. A task that could require several hours can be performed within a few minutes. A widespread adoption of such AI capabilities may disrupt the industry. As outlined in the review of the existing tools, the developers must include such capabilities not only in their interest to compete with others but also in the interest of the whole industry, where severe accidents occur daily.

7 Conclusions

In conclusion, the development and preliminary evaluation of the SWMS AI tool demonstrates its potential to revolutionize safety risk assessments in the construction industry, particularly for small to medium-sized enterprises (SMEs) that often lack sufficient safety resources. By leveraging RAG and agentic AI, the SWMS AI platform can efficiently generate comprehensive and accurate Safe Work Method Statements (SWMS). This underpins the potential contribution of RAG as it is capable of efficiently browsing extensive knowledge bases and retrieving relevant context accurately and reliably.

The examples provided in this work illustrate the tool's capability to identify relevant hazards and suggest appropriate control measures across various construction tasks. The qualitative analysis indicates high accuracy in the generated information, although the repetition of suggested controls in some cases highlights the need for more concise documentation to improve usability.

Future research should focus on quantitatively assessing the system's reliability, accuracy, and usability in real-world settings. Additionally, integrating the SWMS AI tool with other construction safety tasks, such as on-site safety inspections and virtual safety training, could further complement the implementation of RAG-enabled systems in construction safety. The potential inclusion of this tool in digital twin environments for construction safety offers a promising avenue for providing structured, real-time feedback and knowledge to end users, ultimately contributing to reducing workplace hazards.

Acknowledgments

The research presented in this paper was in part funded by the European Union Horizon 2020 research and innovation program under grant agreement no. 101058548.

References

1. Pinto A (2014). "QRAM a qualitative occupational safety risk assessment model for the construction industry that incorporate uncertainties by the use of fuzzy sets." Safety Science, 63. https://doi.org/10.1016/j.ssci.2013.10.019
2. Bureau of Labor Statistics (2022). National census of fatal occupational injuries in 2021. https://www.bls.gov/news.release/archives/cfoi_12162022.pdf (Last accessed, July 31, 2024)
3. Eurostat (2023). Fatal Accidents at work by NACE Rev. 2 activity. https://ec.europa.eu/eurostat/databrowser/view/HSW_N2_02/default/table (Last accessed, July 31, 2024)
4. FIEC EU (2023). Labour shortages in the construction industry: a growing issue. https://www.fiec.eu/application/files/6216/9711/5277/2023-10-12_Shortage_of_Labour_in_the_Construction_industry.pdf (Last accessed, July 31, 2024)
5. Petroni F. et al. (2019). "Language models as knowledge bases?". EMNLP-IJCNLP. https://doi.org/10.48550/arXiv.1909.01066
6. Massarelli L. et al. (2020). "How decoding strategies affect the verifiability of generated text." Findings of the Association for Computational Linguistics Findings of ACL: EMNLP 2020. https://doi.org/10.48550/arXiv.1911.03587
7. Rabbi A.B.K., Jeelani I. (2024). "AI integration in construction safety: Current state, challenges, and future opportunities in text, vision, and audio based applications." Aut Con, 164. https://doi.org/10.1016/J.AUTCON.2024.105443
8. Hussain R. et al. (2024). "Conversational AI-based VR system to improve construction safety training of migrant workers." Automation Construction, 160. https://doi.org/10.1016/j.autcon.2024.105315
9. Lewis P. et al. (2020). "Retrieval-augmented generation for knowledge-intensive NLP tasks." Advances in Neural Information Processing Systems. https://doi.org/10.48550/arXiv.2005.11401
10. SWMS AI (2024). AI-based platform and application for creating safe work method statements. https://www.app.swms.ai/ (Last accessed July 31, 2024).
11. Bhagdev R. et al. (2008). "Hybrid Search: Effectively Combining Keywords and Semantic Searches." http://dx.doi.org/10.1007/978-3-540-68234-9_41
12. Hershcovich D., Donatelli L. (2021). "It's the Meaning That Counts: The State of the Art in NLP and Semantics." KI - Kunstliche Intelligenz 35:255–270. https://doi.org/10.1007/s13218-021-00726-6
13. Huang Q. et al. (2024) "Position Paper: Agent AI Towards a Holistic Intelligence." https://doi.org/10.48550/arXiv.2403.00833
14. Sawarkar K. et al. (2024). "Blended RAG: Improving RAG (Retriever-Augmented Generation) Accuracy with Semantic Search and Hybrid Query-Based Retrievers". https://doi.org/10.48550/arXiv.2404.07220
15. OSHA (2024). "Hazard Prevention and Control: What Is the Hierarchy of Controls?". https://www.osha.gov/safety-management/hazard-prevention (Last accessed, July 31, 2024)
16. Speiser K., Hong K., Teizer J. (2023). "Enhancing the Realism of Virtual Construction Safety Training: Integration of Real-Time Location Systems for Real-World Hazard Simulations". 23rd International Conference on Construction Applications of Virtual Reality (CONVR). https://doi.org/10.36253/979-12-215-0289-3.15
17. Speiser K., Teizer J. (2024). "Automatic creation of personalised virtual construction safety training using digital twins." Management, Procurement and Law. https://doi.org/10.1680/jmapl.23.00104
18. Teizer, J., Johansen, K.W., Schultz, C.L., Speiser, K., Hong, K., Golovina, O. (2024). "A Digital Twin Model for Advancing Construction Safety." CLEaR 2023, Lecture Notes in Civil Engineering (LNCE 390), ISBN: 978-3-031-44020-5, pp. 201–212. https://doi.org/10.1007/978-3-031-44021-2_22
19. Speiser K., Teizer J. (2023). " An Ontology-Based Data Model to Create Virtual Training Environments for Construction Safety Using BIM and Digital Twins". 30th European Group for Intelligent Computing in Engineering Conference.