Safety Management and Human Factors, Vol. 189, 2025, 13–23
https://doi.org/10.54941/ahfe1006566

Development of an Analysis Support Tool for Near-Miss Events Using AI Technology-Improving Human Factor Management Capabilities in the Field

Yuka Banno and Yusaku Okada

Keio University, Science for open and environmental systems, Yokohama, Japan

ABSTRACT

Near-miss analysis is essential for identifying factors contributing to human errors and developing preventive measures. However, conventional text mining methods primarily extract direct causes, such as "lack of attention" or "insufficient verification," often overlooking broader background factors embedded in work environments. This study explores the use of Large Language Models (LLMs) to enhance factor analysis by capturing a more comprehensive range of underlying causes. Using Llama3-ELYZA-JP-8B, we incorporated 87 predefined background factors based on established frameworks, including PSF lists and the m-SHEL model. The developed factor analysis support system was applied to actual near-miss reports, and its effectiveness was evaluated by comparing the number and diversity of extracted factors before and after implementation. Results showed that LLM-based analysis significantly increased factor extraction and enhanced the identification of diverse causes. Additionally, factor aggregation and visualization improved the interpretation of trends over time. Despite these advantages, challenges remain, particularly regarding biases in data, factor extraction, and decision-making. Future research should focus on managing these biases through data diversity, optimized extraction balance, and improved transparency in analysis. By addressing these issues, a more reliable and practical near-miss factor analysis support system can be developed, contributing to improved workplace safety and more effective error prevention strategies.

Keywords: Near-miss analysis, Large language models, Human error prevention, Factor extraction

INTRODUCTION

Various preventive measures against human errors are increasingly being implemented based on information from near-miss events. However, the process from collecting and analysing near-miss events to formulating countermeasures requires significant effort, and the collected near-miss incidents are not being fully utilized.

The purpose of analyzing the factors of near-miss events is to widely collect cases that could potentially lead to incidents, analyze their causes, and develop improvement measures to prevent incidents. By implementing countermeasures based on the identified factors, the likelihood of errors can be reduced, thereby lowering the risk of accidents. To effectively prevent future human errors, it is crucial to identify not only factors that appeared at the time of trouble or direct causes such as "lack of attention" or "insufficient verification," but also broader factors that may have influenced human errors, including those related to tasks and the working environment, which are always present in the target operations. However, in the current situation, most of the reported factors are attributed to the individuals' lack of verification or carelessness, and insufficient consideration is given to factors that are consistently present in the work and environment.

In conventional factor analysis methods, the following two approaches have been considered for analyzing and visualizing recorded text data:

1. Extracting and displaying keywords
2. Visualizing conceptual relationships and co-occurrences between words

The first approach includes methods such as using normalized word frequency, applying TF-IDF to assign weights to words, and utilizing tools like TRENDREADER. The second approach involves methods such as using ontology, leveraging word co-occurrence information, and employing self-organizing maps.

However, there are several challenges in using text data analysis to support factor analysis, including:

- The quality of the presented perspectives depends on the original recorded information.
- Bias in information due to the lack of data or context.
- Subjectivity in the selection of extracted information.

Recently, natural language processing (NLP) using AI has been applied to the analysis of near-miss events. For example, advanced large language models (LLMs) such as BERT (Bidirectional Encoder Representations from Transformers) and GPT are utilized to support similarity analysis with past cases and root cause analysis (RCA).

Therefore, this study aims to leverage Large Language Models (LLMs) to enhance factor analysis of near-miss events. By utilizing LLMs, we enable the extraction of multifaceted underlying factors that consistently exist in work environments and operations, beyond the direct causes reported in individual trouble incident reports. By simplifying the factor extraction process with LLMs, even individuals with limited expertise in human factors can identify less apparent causes and grasp the broader context of incidents. This approach enhances analysts' awareness and contributes to improving workplace safety.

METHODOLOGY

In previous near-miss analyses, text mining has been applied to near-miss reports to extract underlying factors included in these reports. To extract these background factors, PSF (Performance Shaping Factor) lists have been utilized. PSFs refer to factors that influence human behavior, and various lists have been developed to facilitate the search and selection of relevant factors when human errors occur.

Table 1: Background factor.

m-SHEL Category	Subcategory	Background Factor
Management	Organizational Issues in Teams	Difficult to point out issues due to relationships/culture
		Insufficient workforce
		Lack of experienced/skilled workers
		Unclear role distribution
		Overly strict rules/morals
	Education & Management	No workplace education
		Insufficient workplace education
		No preparation
		Insufficient preparation
		Lack of sharing past cases
		Inadequate corporate management system
Software	Information	Insufficient information
		Excessive information
		Inaccurate information
		Scattered information
	Procedures & Planning	Non-compliance with procedures
		Unclear procedures
		No reference materials
		Insufficient reference materials
		Lack of understanding of overall process and objectives
		No checklist
		Insufficient checklist
Hardware	Equipment	Insufficient equipment displays
		Error-prone equipment design
		Similar equipment
		Specialized equipment
		Lack of equipment maintenance
		Poor equipment layout
		Defective equipment
Environment	Task Characteristics	Performing multiple tasks in parallel
		Working alone
		Task interruptions
		Lack of feedback on work results
		Delayed feedback on work results
		Difficult to correct/intervene in tasks
		Strict precision requirements
	Time	High time pressure
		Shift work
		Night work
		Early morning work
		Overtime work
		Long working hours

m-SHEL Category	Subcategory	Background Factor
Liveware (Other People)	Intergroup Relationships	Insufficient collaboration with other groups/industries
	Communication	No confirmation by others
		Insufficient confirmation by others
		No reporting or communication
		Insufficient reporting or communication
		No work instructions
		Insufficient work instructions
Liveware (Individual)	Personal Traits	Not proactive
		Advanced age
		Fear of failure
		Short-sighted thinking
		Reflexive actions
		Habitual actions
		Lack of morals
		Lack of expertise
		Mismatch with physical/skill conditions
	Physical Burden	Physically demanding tasks
		Restricted working posture
		Difficult to memorize
	Mental Burden	Repetitive tasks
		Monotonous tasks
		High task difficulty
		Low task difficulty
		Difficult identification
	Skills	Difficult interpretation
		Difficult prediction
		Difficult judgment
		Unfamiliar tasks
		Familiar tasks
Environment	Work Environment	Work interruption due to non-work-related issues
		Confined space
		Large workspace
		Unstable work environment
		Working at heights
		Excessive lighting
		Insufficient lighting
		Noise
		Vibration
		Poor layout
		Poor visibility
		Disorganized workspace
		Weather conditions
		No fixed placement for tools
		Insufficient workforce
		Excessive workforce

Examples of PSF lists include the PSF list from THERP, the PSF list from HEART, the PSF list from CREAM, and the GAP-W type PSF reference list by Kōtai and Nagata. Based on such PSF lists, background factors of human errors were collected, and a selection list for these factors was created. Additionally, 14 subcategories were defined based on the m-SHEL model, and the collected background factors were classified and organized accordingly.

By referring to these past factor frameworks and classification systems, a total of 87 background factors were identified, which are summarized in Table 1. In this study, these background factors were incorporated into an LLM to facilitate factor extraction.

In this study, we decided to use Llama3-ELYZA-JP-8B, which was developed by further training Meta's Llama 3 model. Here, classification definitions and classification examples were assigned to each of the 87 background factors. The structure of this tool is summarized in Figure 1.

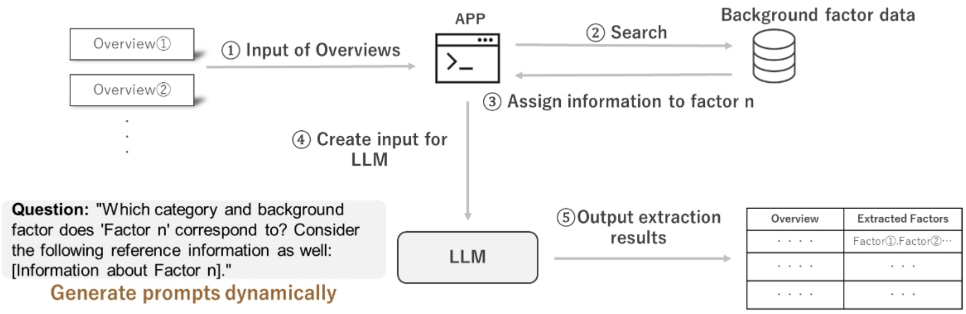


Figure 1: Mechanism of the factor analysis functionality.

There are regular tasks and non-regular tasks in operations. Here, we describe the extraction status of background factors for direct factors that are always present in task execution and environmental conditions during countermeasure operations (Figure 2).

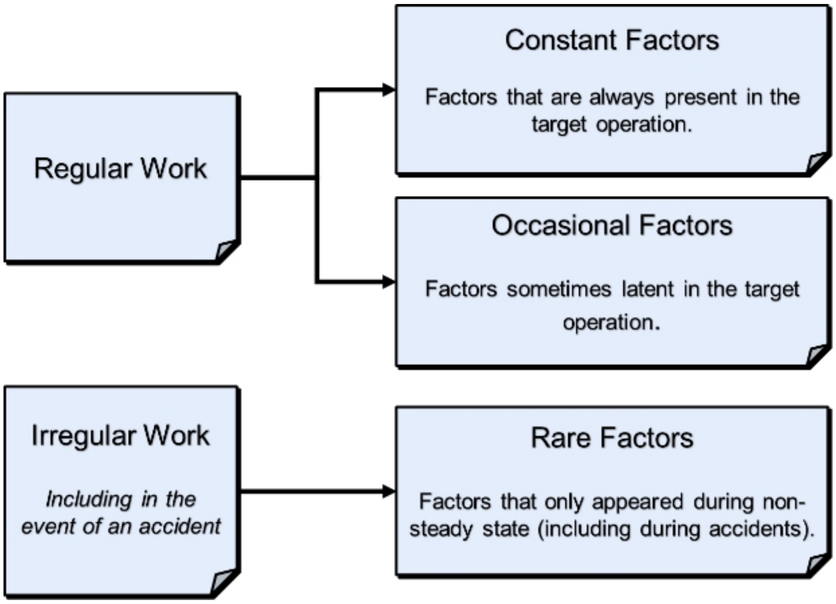


Figure 2: Classification according to the routine nature of the work.

Therefore, documents should collect not only accident-specific information but also factors that affect routine operations. By analysing these factors, they can be utilized for safety measures and recurrence prevention (Figure 3).

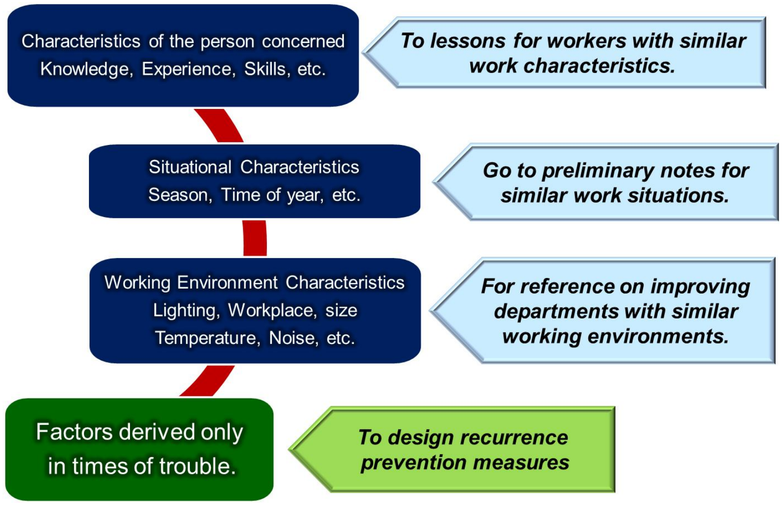


Figure 3: Efforts in information collection.

Furthermore, since an operator's decision-making is influenced by multiple management factors, utilizing the analysis results of near-miss incidents enables more efficient safety management (Figure 4).

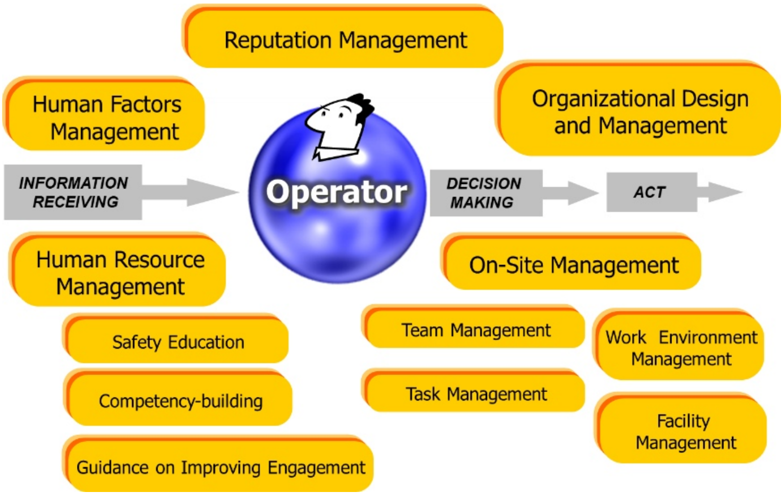


Figure 4: Use of factor analysis for near-miss events for various types of management.

Here, the factor analysis support application developed in this study was applied to the actual analysis work. Factor analysis was conducted for each individual report, and the average number of extracted factors was calculated and compared, as shown in Figure 5. The method used before implementing the factor analysis support application involved extracting background factors through text mining of commonly used report documents.

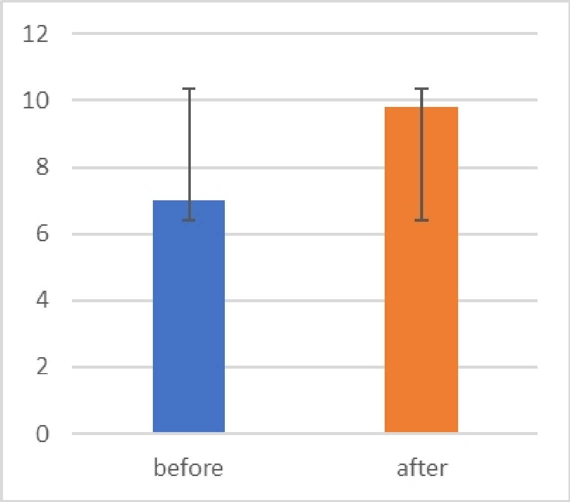
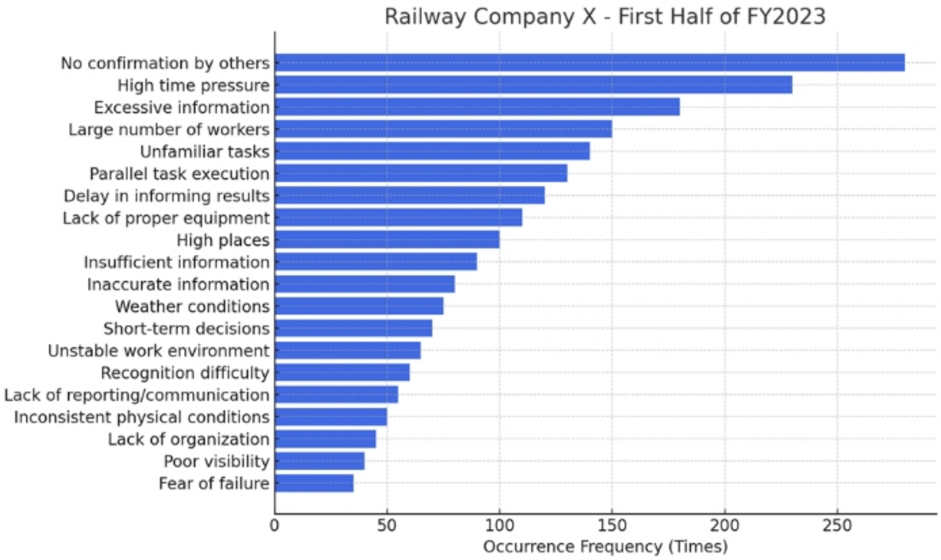


Figure 5: Comparison before and after support.

From Figure 5, it can be observed that when the factor analysis support application developed in this study was applied to actual analysis tasks, the number of extracted factors increased. This indicates that potential factors were identified, and the diversification of extracted factors was confirmed.

As a potential application of this system, aggregating the extracted factors, including latent factors, from documents can further support the factor extraction process. In practice, the extracted factors were aggregated using the hazard prediction factor analysis support application at Railway Company X (Figure 6). Here, the extracted background factors were ranked in descending order of frequency of occurrence, and the top 20 were compiled into a graph. This visualization allows for the interpretation of changes over different periods.



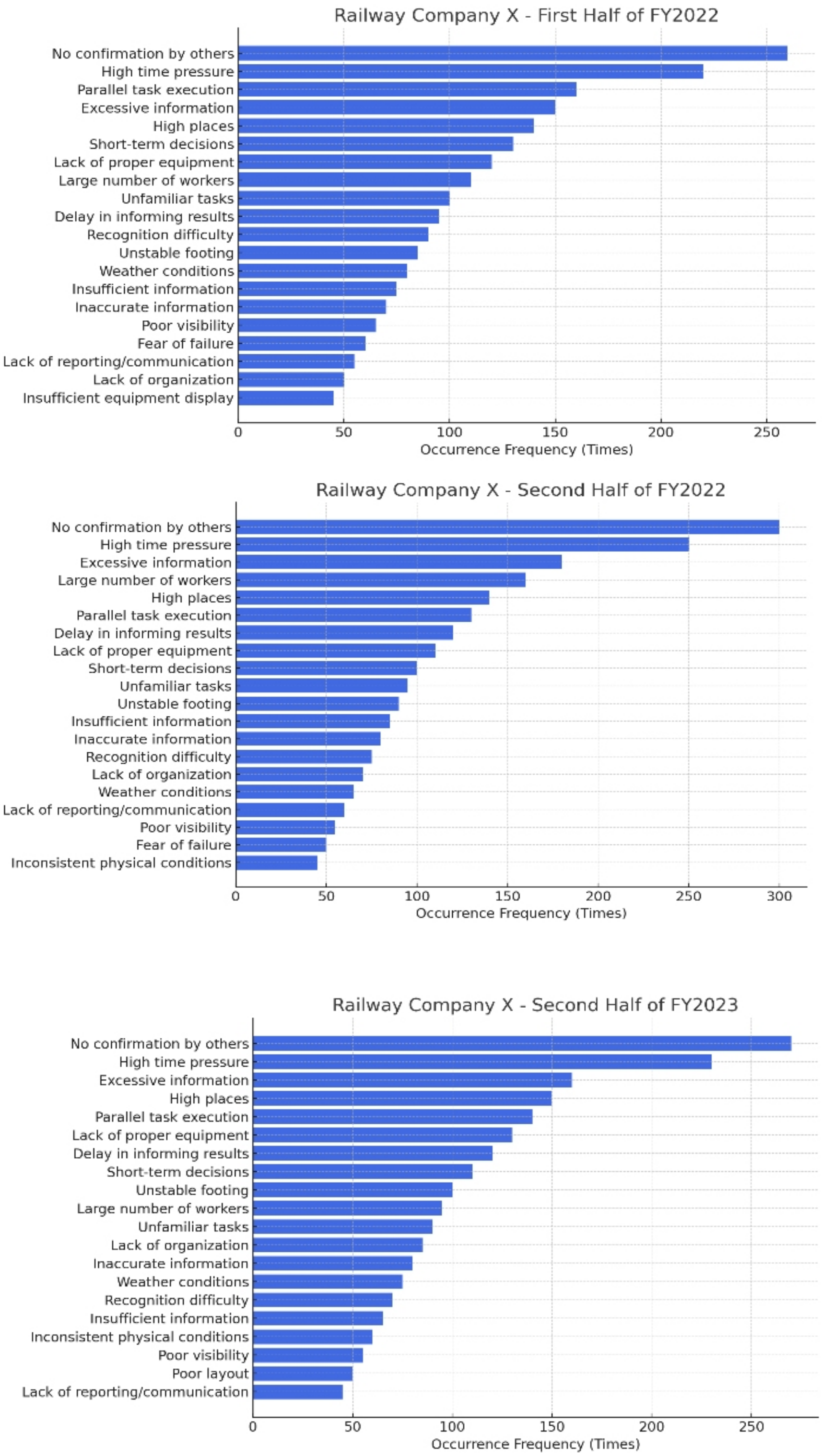


Figure 6: Top 20 occurrences of background factors in company X (FY2023, FY2022).

DISCUSSION

Based on the analysis results from Company X, factors such as "No confirmation by others" and "High time pressure" consistently rank among the top risk factors. This indicates that human error is the primary risk factor, with a strong tendency for workers to proceed with tasks based on individual judgment. Additionally, the persistent presence of risks such as "Lack of confirmation," "Delay in reporting," and "Lack of organization" highlights the need for reinforcing safety awareness as a key challenge.

By conducting such an analysis, the findings contribute to real-world safety activities, including the prioritization of countermeasures for high-risk factors and the development of training programs based on frequently occurring risk factors. Furthermore, for future utilization, sharing accident information with other companies and analyzing risk trends can facilitate the establishment of industry-standard safety guidelines.

The ultimate goal is to develop a factor analysis support system capable of automatically conducting assessments based on guidelines similar to those presented in Table 2. Therefore, Figure 6 has also been evaluated using this table, demonstrating the feasibility of extracting and evaluating risk factors through a structured factor analysis system.

Table 2: Reference guidelines on factor deepening.

Evaluation of factor extraction	Description on Extracted Factors	Proposed Countermeasures	Features	Ripple Effects and Permeability
Bad	Inadvertently Dimly Mistake Misreading Mishearing) Forgetting	Strengthening awareness Saying you will do your best.	Simply reinforcing awareness, but this is something the person concerned is fully aware of on reflection. Being told this several by several people only increases the mental burden on the person concerned and, conversely, increases the likelihood of error.	× It has absolutely no effect on anyone other than the error party.
	Carelessness Assumptions	Reinforced Attention Encouraging Confirmation.	This makes sense if the work was 'no checking action at all' or 'not paying attention', but if it is not, it is just as ineffective as 'I'll do my best' measures, unless it is examined in the context of the current work situation, including the checking work.	× It has absolutely no effect on anyone other than the error party.
Insufficient	Lack of Skills Insufficient experience	Implementation and Thoroughness of Education	This makes sense if there was no education on the subject work, but if not, the instructor may just be stricter, which would not be effective at all.	△ Difficult to obtain visible improvements
	Insufficient documentation Inadequate manuals Inadequate instructions Insufficient Communication	Review of Materials Review of Manuals Clear Instructions Thorough Reporting and Communication	The lack of clarity on how it can be reviewed and how it can be considered clear, the guidelines and goals of the measures, is likely to end up being a formality.	△ The broad outlines of the problems to be addressed could be disseminated but are unlikely to work.
A little Good	Specific Problems in the Operation	Setting up Rules Setting up Prohibitions	Care needs to be taken to ensure that the rules do not become a formality. This should be linked to a survey of employee awareness of the measures.	○ Sustainability is important.
Excellent	Specific and Detailed Problems Involved in the Work	Work improvement Operational improvements Environmental improvements Environmental improvement System improvement	The measures themselves are specific, but they are specific to a certain part of the country; eliminating extractive cause bias with reference to existing factor category methods such as SHEL, 4M, etc.	◎ High spillover effects, but measures to strengthen horizontal development are needed.
	Specific and Detailed Problems Related to the Work	Cross-departmental measures, measures that improve employee satisfaction.	It is desirable to associate this with education and awareness-raising activities for safety activities in general, not just implementation of measures and guidance, so that not only awareness of safety but also understanding and acceptance of safety measures can be achieved.	◎ High spillover effects, but measures to strengthen horizontal development and to develop human resources are needed.

Visualization played a critical role in supporting decision-making and trend analysis. By ranking the frequency of extracted background factors and presenting them in graphical form, stakeholders could quickly grasp which factors were most prevalent and how their frequency changed over time. This enabled more informed prioritization of countermeasures and resource allocation, particularly in identifying persistent risks or emerging patterns that might otherwise remain unnoticed in raw text data.

CONCLUSION

This study explored the use of Large Language Models (LLMs) to enhance the factor analysis of near-miss events. By leveraging LLMs, the analysis process was improved, allowing for the extraction of not only direct causes but also broader background factors that consistently exist in work environments and operations.

The key findings of this research are as follows:

1. **Increased Factor Extraction:** The application of the factor analysis support system resulted in a higher number of extracted factors compared to conventional text mining methods, confirming the identification of potential and latent factors.
2. **Diversification of Extracted Factors:** The system facilitated the discovery of a wider range of background factors beyond commonly reported direct causes, contributing to a more comprehensive understanding of human errors.
3. **Improved Factor Aggregation and Visualization:** The aggregation of extracted factors and their ranking by frequency enabled a clearer interpretation of trends and changes over different time periods.

By integrating LLM-based analysis, even individuals with limited expertise in human factors can effectively identify underlying causes of near-miss events. This approach enhances analysts' awareness and supports the development of more effective preventive measures, ultimately improving workplace safety.

The objective of human error countermeasures is not the elimination of human error, but 'to do a good job (for each person on site)'. Meaningful human error countermeasures can only be said to be effective if the products and services produced by each individual's work are evaluated as valuable by users and society. This must be a common understanding throughout the company, and the extraction and analysis of factors in near-miss events must not be the sole activity of the safety management department. Human error countermeasures are support measures for on-site work. Good countermeasures can only be developed if they are liked and relied upon by the frontline.

We intend to implement a function that can provide specific suggestions on what kind of guidance should be given in the target department regarding the writing of reports, as well as a function that can provide suggestions for modifications to safety guidance in the field, including OJT, so that we can contribute to the development of human error prevention activities.

REFERENCES

Barnett, S., Kurniawan, S., Thudumu, S., Brannelly, Z., Abdelrazek, M. (2024). "Seven Failure Points When Engineering a Retrieval Augmented Generation System." *arXiv preprint*, arXiv:2401.05856.

Chandler, F. T., Chang, Y. H. J., Mosleh, A., Marble, J. L., Boring, R. L., and Gertman, D. I. (2006). *Human reliability analysis methods: selection guidance for NASA*, NASA/OSMA Technical Report, July 2006.

Chew, T. S. (2011). "The checklist manifesto: how to get things right." *Clin Med (Lond)*, 11(3): 296–297.

Dhillon, B. S. (2019). *Safety, reliability, human factors, and human error in nuclear power plants*. CRC Press.

Embrey, D. E., Humphreys, P. C., Rosa, E. A., Kirwan, B., Rea, K. (1984). *SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Structured Expert Judgment*, Volume 1, p. 13, U. S. Nuclear Regulatory Commission, Washington DC.

ELYZA. "Llama3-ELYZA-JP-8B," *Hugging Face*, https://huggingface.co/elyza/ELYZA-japanese-Llama-2-7b.

Gyomachi, Takeo, Nagata, Manabu. (2004). "A Reference List and Case Study Using the GAP-W Type PSF." *Human Factors*, Vol. 9, No. 1, pp. 46–62.

Hawkins, F. H. (1987). *Human Factors in Flight*. Gower Technical Press.

Hollnagel, E. (1998). *Cognitive Reliability and Error Analysis Method (CREAM)*. Elsevier.

Hollnagel, E. (2017). *Safety-II in Practice: Developing the Resilience Potentials (1st ed.)*. Routledge.

Japan Aerospace Exploration Agency (JAXA). (2017). *Human Factors Analysis Handbook*.

Kawamura, R., Inoue, M. (2008). "Development and Application of J-HPES." *Human Factors and Ergonomics Journal*, 50(5): 491–498.

Konishi, Tadawo, Okada, Yusaku. (2002). "A Study on Human Error Factor Analysis Support." *Journal of the Japan Society for Plant Human Factors*, Vol. 7, No. 1.

Llama3. https://www.llama.com/

Miyaji, Yumeko. (2017). "Proposal for the Collection and Utilization of Risk-Related Information such as Near-Misses in Railways." *REAJ Journal*, Vol. 39, No. 6 (Issue 238).

Nakanishi, Miwa, Okada, Yusaku. (2005). "Proposal for an Unsafe Event Collection System Including Near-Misses in Open Mechanisms—Targeting University Chemistry Laboratories." *Human Factors*, Vol. 9, No. 2.

Okada, Yusaku. (2004). "Development of a Human Error Prevention Approach—Analysis of Human Error Seeds and Evaluation of Human Error Probability." *Safety Engineering*, Vol. 43, No. 3.

Read, Gemma J. M. (2017). *Integrating human factors methods and systems thinking for transport analysis and design*. CRC Press.

Reason, J. (2000). "Human error: models and management." *BMJ*, 320(7237): 768–770.

Salmon, Paul M. (2011). *Human factors methods and accident analysis: practical guidance and case study applications*. Ashgate.

Stanton, Neville, Salmon, Paul M., Rafferty, Laura A., Walker, Guy H., Baber, Chris, Jenkins, Daniel P. (2013). *Human factors methods: A practical guide for engineering and design*. Ashgate.

Suzuki, Daisuke, Okada, Yusaku. (2008). "A Study on Education and Training Based on Human Behavior Models—A Proposal for Improving Safety in Train Operations." *Safety Engineering*, Vol. 47, No. 3.

Swain, A. D., Guttmann, H. G. (1983). *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications - Final Report*, NUREG/CR-1278, U. S. Nuclear Regulatory Commission.

Vesely, W. E., Goldberg, F. F., Roberts, N. H., Haasl, D. F. (1981). *Fault Tree Handbook*. U. S. Nuclear Regulatory Commission.