Building AI-Assisted Risk Scoring Models for Proactive Safety Management

Emmanuel Ijeoma
American Society of Safety Professionals (ASSP), US
Institute of Safety Professionals of Nigeria (ISPON), Nigeria
Nigerian Institute of Management (NIM), Nigeria
National Safety Council (NSC), US.

ABSTRACT

For high risk sectors, the typical high impact incident is not caused by a single catastrophic event. More frequently, they are the result of "noise" in the signals that exist in everyday work, but are not strong enough or organized enough to be detected. Traditional safety systems focus on reporting incidents, audit, inspection and lagging indicators to identify an event once it has happened. These methods, while effective in enhancing organisation learning, are constrained by human recognition, inconsistent data interpretation and post hoc analysis. This type of situation introduces a critical time lag between the risk itself and when it becomes visible in the business. This paper discusses evolution and implementation of Artificial Intelligence (AI) aided risk scoring systems for proactive hazard identification and safety management. Utilising recent developments in machine learning, NLP, computer vision, and predictive analytics, the study investigates the possibilities of operational data sources such as near-miss narratives, Permit-to-Work records, environmental sensor data, inspection logs and behavioural indicators being combined to enable the identification of early patterns of increased risk conditions before incidents occur. The research brings evidence from predictive safety research literature as well as operational knowledge from the oil and gas and pipeline building and manufacturing sectors and considers both the technological and organizational aspects of predictive hazard intelligence. The results show that AI-driven systems have a marked ability to catch very subtle and non-obvious risk patterns that are not noticed through regular safety management routines. The usefulness of predictive systems, however, is demonstrated to be not only dependent on the accuracy of the algorithms, but also on the interpretability, contextual validation, trust, and human supervision. The study shows that the most value predictive intelligence has when being implemented as part of the operational decision making process, for example, Permit-to-Work systems, Go/No-Go assessments and frontline safety reviews.

KEYWORDS

Artificial Intelligence in Safety, Predictive Hazard Intelligence, Risk Scoring Models, Safety Systems Integration, Machine Learning, Human-AI Interaction, High-Operational-Risk Industries, Proactive Safety, Operational Risk Prediction

I. INTRODUCTION

A. *Evolution of Safety Management*

Development of safety management has historically been characterized by a predictable pattern of sequential stages: reactive response to proactive prevention. Early industrial systems were based on investigation, learning after failure had taken place. Organizations have since developed and adopted leading indicators such as near-miss reporting, safety audits, behavioural observations, and others to predict and preclude incidents (Reason, 2000). However, a fundamental limitation is still present: the majority of safety systems are dependent on humans to recognize signals of risk. This human recognition brings structural limitations to complex and hazardous systems. Attention spans, reporting, and weak signals are filtered out as background noise. Thus, risk is often recognised only when it has escalated to a point where action is required. This is especially true in industries like oil and gas, construction, and manufacturing, where there is significant variability in the operation and hazard conditions rapidly change.

A new approach to this problem may be provided through the advances of artificial intelligence (AI). Through their capacity to analyse large volumes of structured and unstructured data, machine learning models are proving increasingly successful at detecting patterns that may not be apparent. Research suggests predictive models based on AI can enhance the accuracy of hazard detection and support early detection of risk factors from near-miss data, sensor readings, and log data.

B. Challenges in AI-Driven Safety Systems

However, integrating AI into safety management poses fundamental questions that go beyond the question of technological feasibility. Although predictive accuracy has increased, concerns regarding data quality, model bias, interpretability, and overreliance on (automated) systems continue to exist. The challenge isn't just to develop models that do a good job predicting risk, but to make sure those predictions are valid contextual predictions, practically useful, and ethically managed.

This research takes up this challenge by exploring the design, development, testing, and deployment of risk scoring models supported by AI in hazard identification and control strategies. It goes beyond the usual debates on predictive analytics in safety to consider a less explored aspect: the relationship between human knowledge and computer-intelligence in high-reliability settings.

II. WHEN DATA SEES WHAT HUMANS DON'T

A. The Problem of Fragmented Data

A key weakness in conventional safety management has not been the lack of data, but its inadequate understanding. In today's industrial settings, a plethora of data is available - inspection reports, permit-to-work forms, sensor readings, incident reports, and free-form accounts of near misses. However, these information sources are often treated in isolation, and are analysed within soloed functional areas rather than as elements of a risk system.

This is where artificial intelligence comes in, facilitating data integration and pattern matching. Machine learning models can detect relationships between different data sets that may not be obvious through traditional data analysis. For instance, research indicates that integrating near-miss narratives and environmental sensor data into early warning models for detecting hazard increases their predictive performance, as words in reports tend to have patterns that provide early clues to risk escalation.

B. Role of Natural Language Processing (NLP)

A key component in this is natural language processing (NLP). Near-miss reports, which have long been considered anecdotal and unreliable, can be exploited as predictive instruments with NLP. Linguistic changes (e.g., an increase in references about "pressure," "delay," or "uncertainty") may reflect states of operations under distress that can contribute to an accident. This makes qualitative information into quantitative risk metrics.

Likewise, computer vision technologies have improved in identifying unsafe conditions. Visual AI systems can detect unsafe conditions, like incorrect PPE use, unsafe distances, and equipment malfunctions with growing accuracy, in some cases with over 70% detection in the dynamic environment. These technologies, when combined with predictive analytics, not only detect but also predict hazards by recognising behaviours that lead to incidents.

C. Toward Integrated Predictive Systems

True innovation lies in integrating multiple AI techniques, machine learning, NLP, and sensor analytics, into unified systems. Predictive hazard awareness is achieved when ML/NLP/sensor analytics may be combined and operate as a closed loop system, adapting in real-time based on new input data and continuously refining risk predictions. The literature suggests that such a combined system tends to provide substantially better predictive performances than single models, especially when learning from heterogeneous and high-quality data sets.

But this presents a conundrum. Increasing predictive performance is accompanied by increasing complexity and decreasing explain ability. Since safety professionals must respond to insights derived from models they may not have developed, several questions arise. This brings into question issues of trust, responsibility, and authority. The take-home message is that predictive hazard intelligence is more than just a technological innovation. It is a shift in risk perception, understanding, and assessment - from the visible to the invisible, from human to human-machine interaction.

III. MAKING A PREDICTION IS NOT MAKING A JUDGMENT: THE BIG SAFETY PROBLEM WITH "SMART" SYSTEMS

A. The Limits of Accuracy as a Metric

Most debates about AI and safety end with accuracy. They're wrong. Because the real question is not can AI predict risk - it can. A few studies confirm that ML models, in particular neural networks and ensemble methods, were found to offer the best predictive accuracy when identifying sequences of accident precursors and near-misses. Under certain controlled conditions, models have demonstrated predictive power superior to traditional statistical approaches, especially when trained on large sets.

Here's the problem: What happens when the system is right, but the human does not understand why?

Here's the elephant in the room.

AI-based risk scores can be seen as "narrative-free". A job is considered risky. A location is identified as destabilised. A crew is classified as vulnerable. But the rationale for these markers is embedded among the many layers of algorithmic processing - weights, parameters, and correlations that are difficult to understand.

This introduces a new type of risk: epistemic risk, risk of making a decision based on knowledge that is not explicable.

This is no small matter in safety-critical scenarios. Whether to halt operations, re-process, or make operational interventions has major economic implications. Unless safety practitioners can understand the reasoning behind AI predictions, they could either trust AI or not trust it. Both are risky. This is reflected in the research literature on AI safety systems. Badhan and Samsami (2025) note that false positives and false negatives continue to be a problem, especially in dynamic settings where there is a lot of variability in context. Likewise, Tripathy and Nanda (2025) point out that the quality of the model is highly dependent on data and model calibration in the specific domain of application, and generic solutions with AI would not be reliable if not validated locally.

B. Redefining the Role of Safety Professionals

This is where the safety professional comes in. The safety professional is no longer simply a risk manager, but a risk translator, who needs to cross-check, interpret, and critique the insights generated. This demands a different skill set, blending expertise, data literacy, and critical thinking skills.

So rather than AI replacing human judgment, the future of predictive hazard intelligence is about redefining it. AI can identify patterns. It can detect signals. It can calculate probabilities. But it can't comprehend like us. And consequences are crucial in safety.

IV. PUTTING PREDICTIVE HAZARD INTELLIGENCE SYSTEMS TO WORK

A. Research Methodology and Approach

From conception to implementation, predictive hazard intelligence systems are not simply about algorithms, but the careful design of a system that integrates the structure of information, algorithmic logic, and human decision-making. Our research takes a blended approach of quantitative modelling of production and safety data sets, along with qualitative insights from operational data and experience out of the oil and gas and manufacturing sectors. This is not just to assess predictive performance but also to study the behaviour of such systems in an operational context.

The quantitative part of the research uses multi-source data sets such as near misses, permit-to-work (PTW) logs, inspection, and sensor data on environmental conditions. Machine learning algorithms based on regression (such as random forests and gradient boosting) are used to understand the relationships between leading indicators and incident occurrence. Similar to other studies, these models indicate that one data source is not sufficient; prediction accuracy increases with multiple data sources. For example, the number of near misses has very little predictive power alone, but is a strong predictor when combined with workload, environmental conditions, and complexity of task.

B. Enhancing Prediction through NLP

Linguistic analysis via natural language processing also augments this ability. In this work, NLP models on near-miss narratives show that language features - such as expressions of doubt, time pressure, and procedure violations - are highly correlated with risk states. This confirms previous research that found that when analysed correctly, qualitative data is better than the usual lagging indicators for prediction tasks.

But more important is not how the models perform, how they behave in production. In effective operational settings, predictive information doesn't provide certain conclusions but rather provokes action. Safety professionals use AI risk scores as prompts to launch investigations, verify field conditions, and test hypotheses. This forms a closed loop, which enhances model efficacy.

C. Failure Modes in AI Deployment

On the other hand, the lower-performance systems have two different failure modes. Others are then placed on a pedestal, which results in less critical thinking and slower reaction to model errors. Such results are consistent with the overall human-AI interaction literature emphasizing system interpretability and user engagement as key factors to effective system adoption.

In this light, the paper presents a layered validation schema for risk scoring based on AI. Models are statistically validated at level one with historical data. Outputs are validated operationally at level two through pilot deployments and scenario-based analysis. Level three is continuous validation with user feedback and learning from incidents. By operating on multiple layers, such predictive systems stay dynamic, context-sensitive, and in tune with the real world.

The findings also point to a key organisational need: predictive intelligence must be integrated into operational processes, not exist as a stand-alone entity. Linkage to the PTW process, Go/No-Go decision making, and safety meetings ensures that predictive intelligence can be used. Otherwise, all the models in the world won't make a difference. In conclusion, predictive hazard intelligence is more about a system perspective than technological complexity - coherence of data, models, processes, and people to ensure a coherent approach to safety.

V. CONCLUSION

A. From Reactive Safety to Predictive Intelligence

The drive towards zero harm in high-hazard industries has long relied on the identification of hazards prior to incidents. Predictive hazard intelligence is an important advancement in this regard, moving away from failure identification and towards condition prediction. This research shows that artificial intelligence can improve hazard detection by identifying subtle and dispersed signals in complex systems. The technologies of machine learning, natural language processing, and computer vision allow a greater degree of pattern recognition than conventional human-centred approaches. But the real benefit of these technologies is not the accuracy of the predictions, but in how they are used.

B. Transformation of Human Roles

Our research findings confirm that predictive systems do not take over human expertise - they transform it. Safety practitioners shift from playing the key role of risk identification to interpreting and validating decisions made by machines. This raises new issues of trust, transparency, and accountability, including specific governance and validation processes.

Importantly, the research shows that the impact of predictive hazard intelligence is dependent on design. Decentralised deployments, in which AI is used outside of operational processes, have little effect. On the other hand, connected systems, where predictive insights feed into the planning, execution, and review of operations, result in increased safety and stability.

The take-home message for industry is two-fold. Companies that want to implement smart safety systems based on AI technologies need to invest not only in technology but also in their people, data, and cultural alignment. Predictive intelligence is not something to be used, but rather created. Ultimately, safety will be defined not by the speed of organisations' reactions, but by their ability to see the potential for accidents to occur. AI offers the opportunity to see them sooner. It will be up to organizations to see.

C. Researcher's Bio

Emmanuel Ijeoma is a well-trained Health, Safety and Environmental (HSE) consultant.

REFERENCE

[1] Akash, M. A. S. (2026). Artificial intelligence enabled predictive safety analytics for proactive hazard identification in industrial manufacturing environments. International Journal of Scientific Interdisciplinary Research, 7(1), 416–452. https://ijsir.org/index.php/IJSIR/article/view/89

[2] Akwaisua, A. M., Ekong, A., & Ansa, G. (2025). Enhancing hazard detection and risk severity assessment in construction through multinomial naive Bayes and regression. Journal of Information Systems and Informatics, 7(1), 97–121. http://journal-isi.org.adsii.or.id/index.php/isi/article/view/979

[3] Siqueira, M., Martins dos Santos, V., Diniz, E., & Paula, A. (2024). Artificial intelligence for sustainability: A systematic literature review in information systems. Revista de Gestão Social e Ambiental, 18, 1–20. https://doi.org/10.24857/rgsa.v18n3-178

[4] Badhan, S. J., & Samsami, R. (2025). Artificial intelligence in construction safety: A systematic literature review. Buildings, 15(22), 4084. https://www.researchgate.net/profile/Sharmin-Jahan-Badhan/publication/397606034_Artificial_Intelligence_AI_in_Construction_Safety_A_Systematic_Literature_Review/links/693079ec27359023a008a33f/Artificial-Intelligence-AI-in-Construction-Safety-A-Systematic-Literature-Review.pdf

[5] Gnoni, M., & Tornese, F. (2025). Designing a near-miss management system. In ... https://doi.org/10.1201/9781003428947-16

[6] Nadler, D. (2024). Machine learning in occupational health and safety: A review of knowledge gaps. Preprints. https://www.preprints.org/frontend/manuscript/fc9339867ac301d44e0021e5732dd8e0/download_pub

[7] La Torre, G., et al. (2026). Artificial intelligence and occupational health and safety: A systematic review. Journal of Public Health, 1–21. https://link.springer.com/article/10.1007/s10389-026-02738-8

[8] Sarkar, S., et al. (2019). Application of optimized machine learning techniques for prediction of occupational accidents. Computers & Operations Research, 106, 210–224. https://www.sciencedirect.com/science/article/pii/S0305054818300601

[9] Tripathi, S., et al. (2025). Transforming surgical planning and procedures through the synergistic use of additive manufacturing, advanced materials and artificial intelligence: Challenges and opportunities. Materials Horizons, 12(19), 7814–7864. https://pubs.rsc.org/en/content/articlehtml/2025/mh/d5mh00501a