

 ISSN (O): 2320-5407 ISSN (P): 3107-4928	Journal Homepage: www.journalijar.com INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR) Article DOI:10.21474/IJAR01/24377 DOI URL: http://dx.doi.org/10.21474/IJAR01/24377	
--	--	---

RESEARCH ARTICLE

AI-ENABLED PROACTIVE SAFETY INTERVENTION IN OIL & GAS AND CONSTRUCTION PROJECTS: A HUMAN-SUPERVISED PREDICTIVE RISK REVIEW

Rafiullah Khan

Manuscript Info

Manuscript History

Received: 12 August 2026

Final Accepted: 14 September 2026

Published: October 2026

Abstract

Artificial intelligence is rapidly entering construction and oil and gas safety management through computer vision, predictive modelling, natural-language processing, digital twins and sensor analytics. Yet predictive accuracy is not equivalent to safety improvement. This review examines how AI-enabled information can become a defensible safety intervention rather than a parallel technology program. A structured integrative review of peer-reviewed literature published from 2020 to 2025 was conducted across construction safety, upstream and pipeline operations, asset integrity and digital risk management. The analysis focuses on four questions: what safety states can be predicted reliably, how predictions should be validated, how human judgement and authority should be retained, and what organizational conditions determine whether warnings produce risk reduction. Evidence is strongest for visual hazard recognition, injury-pattern classification, PPE and proximity monitoring, predictive maintenance, pipeline integrity and data-supported risk estimation. Important limitations persist in data quality, class imbalance, explainability, model drift, sensor failure, transfer across sites, privacy and automation bias. The review proposes a human-supervised predictive safety intelligence framework linking data acquisition, contextualization, model inference, risk translation, competent validation, controlled intervention and feedback learning. Two design principles dominate: AI should target a clearly defined control problem, and every prediction should connect to an accountable decision pathway. The paper concludes that AI can strengthen proactive safety when it increases the timeliness and reliability of risk information while preserving engineering controls, professional judgement and auditable governance.

"© 2026 by the Author(s). Published by IJAR under CC BY 4.0. Unrestricted use allowed with credit to the author."

Introduction:-

The digitalization of construction and oil and gas operations has created an unprecedented volume of safety-relevant data. Cameras observe work fronts continuously, wearable devices capture movement and physiology, equipment sensors report operating conditions, electronic permits record work states, and asset-management systems contain histories of inspection, corrosion, maintenance and incidents. Artificial intelligence offers methods for converting

this information into forecasts, classifications and decision support. The promise is attractive: move from learning after an accident toward identifying deteriorating conditions before harm occurs.

Recent reviews show substantial growth in construction applications. Deep learning is used to detect workers, equipment, personal protective equipment and unsafe spatial relationships; machine learning is applied to accident databases; and text analytics is increasingly used to structure reports and regulatory requirements (Liu et al., 2022; Rabbi & Jeelani, 2024). Digital safety frameworks combine these tools with project information and workflow systems so that hazard recognition can be linked to site decisions (Larbi et al., 2024). In oil and gas, predictive analytics supports asset integrity, pipeline risk estimation, maintenance prioritization and incident analysis (Sattari et al., 2022; Priyanka et al., 2022; Khan et al., 2024). The common ambition is anticipatory control.

However, safety is a demanding domain for AI. Rare catastrophic events produce limited labelled data, operating conditions change, and a model can be statistically accurate while still being operationally unsafe. A false negative may leave a critical hazard undetected; excessive false positives may create alarm fatigue. Models trained on one contractor population, climate or site layout may not transfer to another. Visual systems can be affected by occlusion, lighting and camera position; sensor systems can drift or fail; language models can generate plausible but incorrect interpretations. Most importantly, safety decisions carry accountability that cannot be delegated simply because an algorithm has high validation accuracy.

These problems make the intervention pathway more important than the algorithm alone. The relevant question is not whether AI can detect a hard hat, predict an injury class or estimate corrosion. It is whether the information arrives early enough, is credible enough, and reaches a person or automated safeguard with authority to reduce exposure. This distinction separates predictive safety intelligence from technical experimentation. In a strong system, prediction is one component of a controlled loop that includes data quality, contextual interpretation, human validation, intervention and learning.

The thesis context of technical operations adds further complexity. Commissioning, shutdowns, heavy lifting, testing, energization and work in live facilities are performed under changing configurations and by multi-employer teams. The same sensor reading may mean different things depending on work phase, permit status or temporary barrier condition. AI systems must therefore represent context rather than only patterns. A proximity alert around a mobile crane, for example, has different significance when lifting is active, suspended, or being prepared. A pipeline pressure anomaly has different meaning during stable operation than during controlled start-up.

This review develops a practical framework for AI-enabled safety intervention across construction and oil and gas projects. It does not argue for autonomous control as a default. Instead, it evaluates where predictive models can enhance professional judgement, where engineering barriers must remain primary, and what governance is required to make data-driven intervention reliable and auditable.

Aim and Research Objectives:-

The aim of this study is to critically evaluate contemporary AI-enabled safety technologies and to develop a human-supervised predictive intervention framework for high-risk construction and oil and gas operations. Five objectives are pursued. First, the review maps the main AI modalities used for safety: computer vision, structured machine learning, natural-language processing, sensor analytics and digital-twin-supported prediction. Second, it compares their value across hazard recognition, incident prediction, asset integrity and operational-risk decisions. Third, it identifies technical and human-factors limitations that can undermine safety performance despite high model accuracy. Fourth, it defines governance requirements for validation, escalation, accountability and fallback. Fifth, it proposes a staged maturity pathway through which organizations can progress from reactive reporting to predictive and adaptive safety management without removing competent human oversight.

Review Methodology:-

A structured integrative review was selected because AI safety research spans engineering informatics, construction management, process safety, asset integrity, computer science and occupational safety. The studies use different data types and outcomes, preventing meaningful statistical pooling. The synthesis therefore emphasizes mechanism and deployment: what data are used, what safety state is inferred, what decision the output supports, how the model is validated, and what constraints affect field use. Searches combined terms including artificial intelligence, machine learning, deep learning, computer vision, natural-language processing, predictive safety, construction safety, oil and gas, pipeline integrity, asset integrity, accident prediction, hazard recognition, digital twin, sensor analytics, PPE

detection, risk assessment and predictive maintenance. Scopus, Web of Science, ScienceDirect, IEEE Xplore and publisher databases were used for discovery, supplemented by citation tracing. The final reference set was limited to peer-reviewed publications from 2020-2025 and each entry includes a DOI.

Inclusion required an identifiable safety or integrity application, sufficient methodological description to understand the data-model-decision relationship, and relevance to construction, oil and gas, infrastructure or closely transferable high-risk operations. Studies were excluded when AI was mentioned only conceptually, when the output had no plausible safety intervention, or when the evidence consisted primarily of vendor promotion. Review papers were retained to map technological trends, while empirical studies were used to examine operational performance and limitations.

Quality appraisal considered six dimensions. Data provenance assessed whether the origin, representativeness and labelling of inputs were clear. Validation assessed whether performance was tested on independent or realistic data. Safety relevance examined whether the predicted variable was causally connected to a control decision. Uncertainty considered false positives, false negatives, calibration and confidence. Human integration assessed explainability, workload, trust and escalation. Deployment realism considered sensor reliability, connectivity, cybersecurity, contractor variability and maintenance of the digital system itself.

Thematic synthesis was organized into visual intelligence, predictive analytics, asset-integrity intelligence, language and knowledge systems, human-system integration and governance. No fabricated study-flow counts are reported because a formal exported screening database was not maintained for this manuscript. The methodological priority is traceability of inclusion logic and critical comparison rather than numerical completeness. The final synthesis also distinguishes algorithmic performance from intervention effectiveness: an accurate model is treated as enabling evidence, not as proof that harm has been reduced.

Critical Literature Synthesis:-

Computer vision and visual hazard recognition:-

Computer vision is one of the most mature AI applications in construction safety. Reviews describe systems for detecting PPE, workers, equipment, unsafe proximity, fall exposure and work-zone intrusions (Baduge et al., 2022; Rabbi & Jeelani, 2024). The attraction is practical: construction sites already generate large volumes of visual data, and many hazards have a visible spatial component. Deep-learning models can convert images or video into structured observations at a scale that manual inspection cannot match.

The safety value depends on what the detection represents. Recognizing that a person is not wearing a hard hat is a direct observation of a rule condition. Recognizing that a worker is near moving plant is more context dependent because safe distance varies with speed, direction, load and barriers. Hou et al. (2023) illustrate the move toward risk computation and visualization rather than simple object detection. This is important because risk is relational: it emerges from the configuration of people, equipment, tasks and controls.

Automated reporting can further connect detection with work processes. Wen et al. (2024) integrate deep learning with visual question answering to produce safety reports, illustrating how AI can reduce the gap between observation and documentation. Yet automation introduces a verification challenge. Generated descriptions may be fluent without being fully accurate, and managers may treat machine-generated reports as objective even when camera coverage is incomplete. A defensible system should preserve the original evidence, report model confidence, identify blind zones and require human confirmation before consequential enforcement or work stoppage when uncertainty is material.

Synthetic image generation is emerging as a method for addressing data scarcity. Kim and Yi (2024) use generative models to create hazardous construction scenarios for training deep networks. The approach can improve class balance where dangerous situations are rare, but synthetic realism must be validated. A model may learn artefacts of generated images rather than features of actual hazards. Synthetic data are therefore best treated as augmentation, with field data retained for final testing.

Predictive models from incident and operational data:-

Structured machine-learning models use historical incidents, project attributes and environmental conditions to estimate risk. Alkaissy et al. (2023) demonstrate classification of injury types using a large accident dataset, while Ajayi et al. (2020) apply deep learning to health and safety risk prediction in power infrastructure. These studies

show that non-linear relationships and high-dimensional data can support useful discrimination beyond simple descriptive statistics.

The central limitation is that historical records reflect the reporting system that produced them. Minor incidents are more common than catastrophic events, definitions change, contractor reporting varies, and near misses may be under-recorded. A model can therefore be accurate about the historical database yet misrepresent actual hazard prevalence. Predictive safety should include data-quality indicators alongside model metrics. Missingness, reporting delays, class imbalance and changes in incident coding are not merely technical nuisances; they are potential sources of unsafe confidence.

Causal interpretation is another challenge. Feature importance identifies variables associated with predictions, not necessarily interventions that will reduce risk. If long working hours correlate with incidents, reducing hours may be sensible; if project size appears important, it may simply proxy organizational complexity. Managers should distinguish predictive features from controllable causes. This is particularly important when models are used to prioritize inspections or allocate resources, because an apparently precise score can conceal uncertainty about the mechanism.

Asset integrity and oil and gas applications:-

Oil and gas applications extend predictive safety beyond worker behaviour to the condition of assets and process barriers. Sattari et al. (2022) propose a data-driven AI decision framework for asset integrity management, emphasizing the need to connect analytics with established integrity processes. Priyanka et al. (2022) combine digital-twin concepts with prognostic and machine-learning methods for pipeline risk estimation. Predictive maintenance reviews similarly show how digital twins can integrate operating data and degradation models to support maintenance decisions (van Dinter et al., 2022).

These applications can contribute to safety because loss of integrity is a precursor to release, fire and environmental harm. However, asset predictions often involve uncertainty about degradation rates, sensor coverage and future operating conditions. The safest use is decision support within a defined inspection and engineering regime. A model may identify a segment requiring earlier inspection, but it should not unilaterally extend inspection intervals without engineering review and documented basis.

External corrosion illustrates the need for combined physical and data knowledge. Wasim and Djukic (2022) review mechanisms and predictive prevention for oil and gas pipelines, showing that corrosion emerges from interacting material, environmental and electrochemical conditions. AI can detect patterns across monitoring data, but causal engineering understanding remains necessary for extrapolation. A purely data-driven model may perform poorly after changes in coating, cathodic protection, fluid composition or operating temperature because the data-generating process has changed.

AI-supported incident analysis is another growing area. Khan et al. (2024) apply machine learning to oil and gas accident and incident analysis. Such tools can help identify recurring causal patterns across large textual or coded datasets. Their organizational value depends on whether lessons alter controls. Classification without feedback into design, procedures, competency or governance risks becoming an analytical endpoint rather than an intervention.

4.4 Natural-language processing and knowledge representation

Safety systems contain large volumes of unstructured text: permits, method statements, incident narratives, lessons learned, regulations and procedures. Natural-language processing can extract entities, relations and obligations that are difficult to manage manually. Wang and El-Gohary (2023) demonstrate relation extraction and knowledge-graph representation of construction safety requirements, illustrating a path toward machine-readable compliance knowledge.

The potential intervention is significant. A knowledge system can flag conflicts between a task plan and a requirement, retrieve relevant controls for a hazard, or identify recurring causes in incident narratives. Yet language systems require careful governance because ambiguity is normal in technical documents. Terminology differs among contractors, abbreviations are local, and the meaning of a requirement can depend on context. Human validation should therefore remain mandatory for interpretations that change work authorization or engineering controls.

Large language models add further capability but also new failure modes. They can summarize, draft and retrieve information rapidly, but probabilistic generation can produce statements unsupported by source documents. For safety-critical use, the system should retrieve from controlled documents, show citations or source passages, restrict authoritative answers to approved content and log model versions. The design objective is not conversational fluency; it is traceable assistance.

Sensors, wearables and multimodal fusion:-

AI becomes more valuable when models can combine multiple data streams. Wearable devices provide location, motion and physiological information; environmental sensors provide heat, gas, noise or air-quality data; equipment systems provide speed, load and status. Maali et al. (2024) review existing and emerging construction safety technologies, while broader digital-safety literature emphasizes the potential of connected systems to move from periodic inspection to continuous observation.

Multimodal fusion can resolve ambiguities that a single sensor cannot. A worker's location becomes more meaningful when combined with equipment motion and task status. Methane concentration becomes more actionable when combined with wind, source location and operating state. Andrews et al. (2023) show how machine learning can support gas-sensor calibration for methane monitoring, illustrating the importance of analytics in maintaining measurement quality as well as interpreting measurements.

The challenge is reliability of the whole chain. Battery failure, communication loss, calibration drift or synchronization errors can make an apparently sophisticated system less reliable than its individual components suggest. Safety architecture should therefore include sensor-health monitoring, data-quality flags and graceful degradation. When digital data are unavailable, essential controls must continue through conventional procedures rather than leaving the work dependent on a fragile information service.

Human oversight, trust and automation bias:-

AI changes the cognitive work of safety professionals. Instead of searching for every hazard manually, users may supervise model outputs. This can reduce workload but can also create automation bias, where users accept recommendations because they appear mathematically authoritative. Gazos et al. (2025) argue that AI-enhanced sociotechnical systems need organizational analysis, not only technical validation. Structural vulnerabilities can emerge from unclear responsibility, weak challenge processes or over-reliance on automated advice.

Trust should be calibrated rather than maximized. Users need to understand when a model is reliable, when it is uncertain and when conditions differ from its training environment. Explainability can help, but explanations should be designed for decisions rather than as generic feature rankings. For a safety professional, the useful explanation is often which observations drove the alert, how far the current case lies from known examples, and what control options are available.

Human oversight also requires authority. A supervisor who receives a high-risk prediction but lacks authority to stop work cannot convert intelligence into intervention. Conversely, automatic shutdown may be appropriate only where response time is too short for human action and the logic has been engineered and validated as a protective function. Most AI safety applications reviewed here are better positioned as advisory or supervisory layers that increase awareness and prioritization while established engineering safeguards retain primary control.

Model drift, transferability and lifecycle assurance:-

Construction projects evolve rapidly. Camera viewpoints change, subcontractors mobilize and demobilize, weather changes, work shifts from structure to finishing, and equipment types vary. Oil and gas assets also change through maintenance, debottlenecking, feed composition and aging. A model that performed well during development can therefore degrade silently. Drift monitoring should be treated as a safety requirement, not a data-science convenience.

Lifecycle assurance includes initial validation, controlled deployment, periodic recalibration, change management and retirement. Any significant change to sensors, labels, thresholds or algorithms can alter performance. Version control and traceable approval are essential where outputs influence safety decisions. The same principle applies to cybersecurity: tampered data or unavailable systems can corrupt risk assessment. AI governance must therefore integrate with existing management of change, asset integrity and project assurance rather than operate as an isolated innovation process.

Human-supervised predictive safety intelligence loop

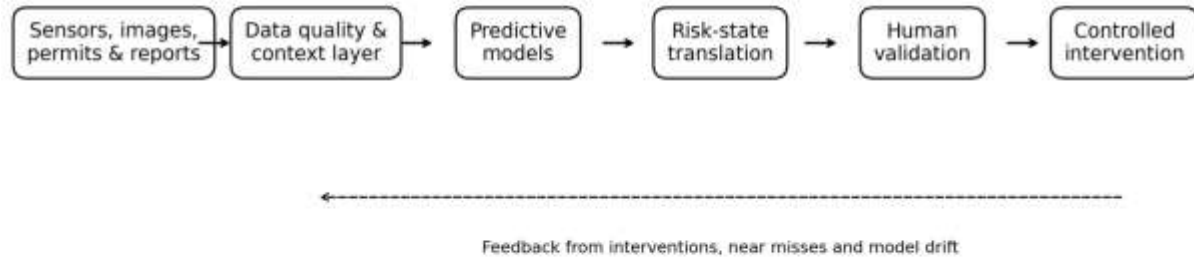


Figure 1. Human-supervised predictive safety intelligence loop from data acquisition to controlled intervention.

Table 1. AI modalities, principal safety applications, strengths and characteristic limitations.

Intervention family	Primary failure mode	Verification evidence	Example leading indicator
Computer vision	PPE, proximity, access, spatial hazards	Broad continuous coverage	Occlusion, lighting, false positives, context ambiguity
Structured machine learning	Incident likelihood, injury type, risk prioritization	Pattern detection across large datasets	Bias, class imbalance, weak causal interpretation
Natural-language processing	Requirements, permits, narratives, lessons learned	Rapid extraction and knowledge linking	Ambiguity, unsupported generation, terminology variation
Sensor analytics	Gas, heat, fatigue, motion, equipment state	Real-time physical-state information	Calibration drift, connectivity, battery and synchronization
Digital twin / hybrid models	Integrity, scenario forecasting, maintenance decisions	Contextual integration of data and engineering models	Model drift, complexity, cybersecurity and maintenance burden

Predictive Safety Intelligence Framework and Discussion:-

The evidence supports a six-stage human-supervised loop. Stage one acquires data from sensors, images, reports, permits, asset systems and environmental sources. Stage two contextualizes and quality-checks those data. Stage three applies predictive or classification models. Stage four translates model output into a risk state that is meaningful for operations. Stage five requires competent human validation unless the function is a pre-engineered automatic protective action. Stage six executes a controlled intervention, after which outcomes and errors return to the data and learning process.

Figure 1 illustrates this loop. Its most important feature is not the model block but the connections around it. Data must be trustworthy, risk categories must correspond to control thresholds, and the person receiving the warning must have defined authority. The framework therefore treats AI as part of a sociotechnical control system. Model accuracy remains important, but it is only one assurance claim among many.

Table 1 summarizes major AI modalities and their strongest safety uses. Visual models are effective for observable spatial hazards; structured machine learning can prioritize risk from historical patterns; language systems can organize safety requirements and narratives; sensor analytics can detect abnormal physical states; digital twins can integrate context and forecast scenarios. Each modality also has characteristic failure modes. This supports a portfolio approach rather than dependence on a single technology.

A maturity pathway is shown in Figure 2. Reactive organizations learn mainly from incidents. Monitored organizations collect continuous data but still rely on thresholds. Predictive organizations use models to anticipate likely deterioration. Adaptive organizations close the loop by learning from interventions, drift and changing conditions. Progression should be evidence based. An organization should not move toward greater automation until data reliability, governance and human integration are demonstrably mature.

The framework also clarifies a common misconception: more prediction does not automatically mean less risk. A warning has value only if it changes exposure or restores a degraded control. Therefore, performance evaluation should include operational metrics such as time from alert to intervention, proportion of high-risk alerts verified, false-alarm burden, missed critical events, recurrence after corrective action and user override patterns. These indicators connect model performance with safety performance.

A predictive model should also communicate uncertainty in a form that supports operational judgement rather than presenting a single score with false precision. Confidence intervals, data-quality flags, and out-of-distribution warnings can help users recognize when model output deserves additional scrutiny. Where consequences are severe, conservative fallback rules should take precedence when required inputs are missing or sensor health cannot be demonstrated. Human factors evaluation should examine whether alerts interrupt work appropriately, whether wording is understandable, and whether repeated notifications create desensitization. Model governance should include a documented route for frontline users to challenge outputs and provide evidence that a prediction was contextually wrong. Those challenges are valuable learning data because they reveal conditions that were absent from development datasets or poorly represented in model features. Organizations should also separate exploratory analytics from production safety functions so that experimental models cannot silently influence critical decisions. Production deployment requires configuration control, access permissions, audit trails, and tested rollback arrangements comparable to other safety-relevant digital systems.

When multiple models are combined, assurance must consider interaction effects because individually acceptable components can create contradictory or unstable recommendations. Data integration should preserve timestamps and provenance, particularly where asynchronous sensors could make two observations appear simultaneous when they were not. Contractor turnover creates another source of drift because work methods, language, PPE appearance, reporting habits, and supervision patterns can change rapidly. Climate and environmental conditions can also alter sensor performance and human exposure, making local validation essential in heat, dust, glare, rain, or low-light conditions. Predictive maintenance models should be aligned with engineering criticality so that resources are not diverted toward easily predicted degradation while high-consequence barriers remain weakly monitored. Safety analytics teams therefore need regular interaction with engineers, supervisors, operators, and workers who understand the physical meaning of the data.

A technically sophisticated model can still fail organizationally when ownership is unclear, warnings are routed to the wrong role, or intervention budgets are unavailable. Performance reviews should examine these organizational bottlenecks with the same seriousness as precision, recall, or area-under-curve statistics. The most mature implementations will likely combine physical models, historical data, expert knowledge, and live sensing rather than relying exclusively on one learning paradigm. Such hybrid approaches can improve interpretability because predictions can be checked against known engineering constraints and expected operating envelopes. Nevertheless, every added data source increases complexity, so architecture should remain proportional to the safety decision and the organization's capacity to maintain it. The goal is dependable anticipation, not technological maximalism, and successful systems will be those whose warnings consistently lead to timely, justified, and auditable control actions. A predictive model should also communicate uncertainty in a form that supports operational judgement rather than presenting a single score with false precision. Confidence intervals, data-quality flags, and out-of-distribution warnings can help users recognize when model output deserves additional scrutiny. Where consequences are severe, conservative fallback rules should take precedence when required inputs are missing or sensor health cannot be demonstrated. Human factors evaluation should examine whether alerts interrupt work appropriately, whether wording is understandable, and whether repeated notifications create desensitization. Model governance should

include a documented route for frontline users to challenge outputs and provide evidence that a prediction was contextually wrong. Those challenges are valuable learning data because they reveal conditions that were absent from development datasets or poorly represented in model features.

Organizations should also separate exploratory analytics from production safety functions so that experimental models cannot silently influence critical decisions. Production deployment requires configuration control, access permissions, audit trails, and tested rollback arrangements comparable to other safety-relevant digital systems. When multiple models are combined, assurance must consider interaction effects because individually acceptable components can create contradictory or unstable recommendations. Data integration should preserve timestamps and provenance, particularly where asynchronous sensors could make two observations appear simultaneous when they were not. Contractor turnover creates another source of drift because work methods, language, PPE appearance, reporting habits, and supervision patterns can change rapidly.

Climate and environmental conditions can also alter sensor performance and human exposure, making local validation essential in heat, dust, glare, rain, or low-light conditions. Predictive maintenance models should be aligned with engineering criticality so that resources are not diverted toward easily predicted degradation while high-consequence barriers remain weakly monitored. Safety analytics teams therefore need regular interaction with engineers, supervisors, operators, and workers who understand the physical meaning of the data. A technically sophisticated model can still fail organizationally when ownership is unclear, warnings are routed to the wrong role, or intervention budgets are unavailable. Performance reviews should examine these organizational bottlenecks with the same seriousness as precision, recall, or area-under-curve statistics. The most mature implementations will likely combine physical models, historical data, expert knowledge, and live sensing rather than relying exclusively on one learning paradigm. Such hybrid approaches can improve interpretability because predictions can be checked against known engineering constraints and expected operating envelopes. Nevertheless, every added data source increases complexity, so architecture should remain proportional to the safety decision and the organization's capacity to maintain it. The goal is dependable anticipation, not technological maximalism, and successful systems will be those whose warnings consistently lead reliably to timely, justified, and auditable control actions.

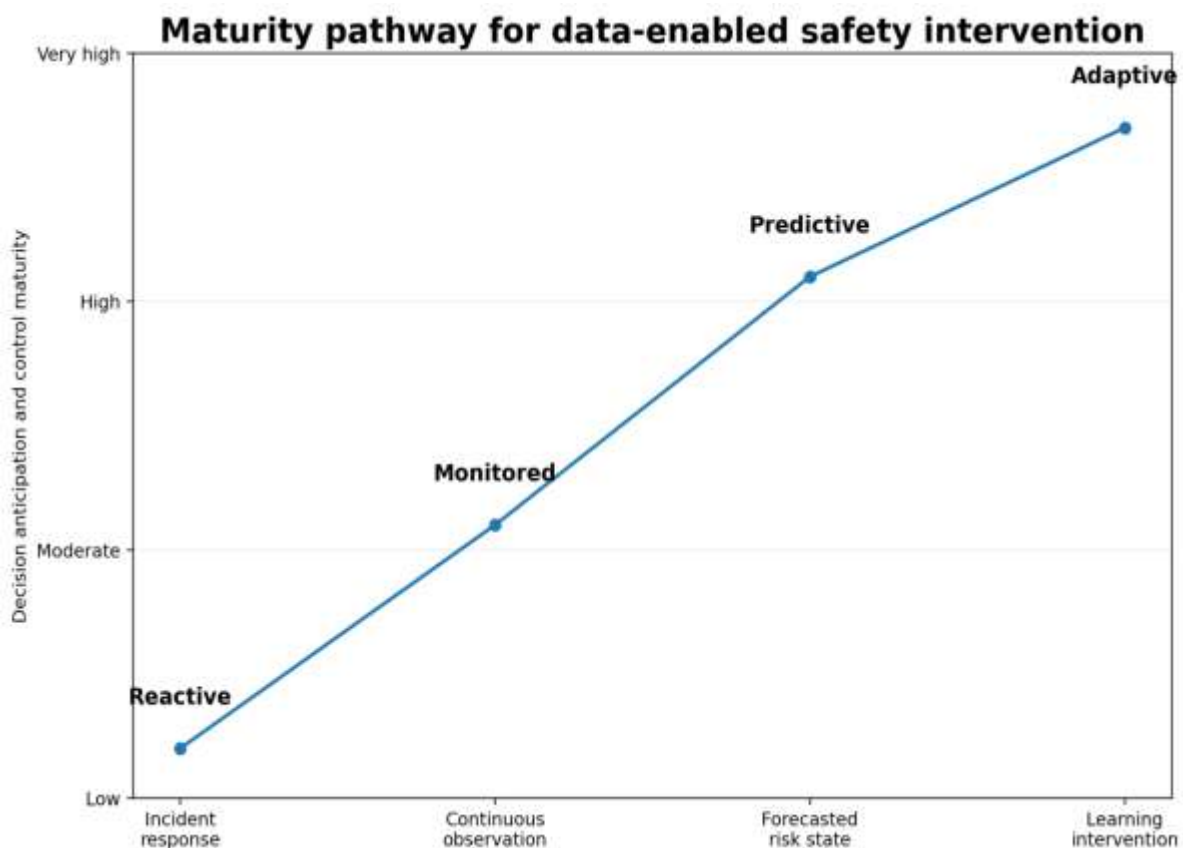


Figure 2. Maturity pathway from reactive response to adaptive, learning-based safety intervention.

Table 2. Assurance criteria for data integrity, model validity, human oversight, control linkage and lifecycle governance.

Assurance domain	Indicator	Interpretation	Management response
Data integrity	Known provenance, calibrated sensors, complete timestamps	Reject or qualify predictions when required inputs are unreliable	Data owner and safety owner
Model validity	Independent testing, calibration, false-negative analysis	Set deployment limits and review thresholds	Model owner with domain reviewer
Human oversight	Explainable alert, competent reviewer, challenge route	Confirm or override prediction with documented rationale	Authorized supervisor or engineer
Control linkage	Prediction mapped to defined intervention and response time	Ensure intelligence changes exposure or restores a barrier	Operational risk owner
Lifecycle governance	Version control, drift monitoring, cybersecurity, fallback	Revalidate after significant change; maintain safe degraded mode	Management of change authority

A comparable discipline is required when predictive systems move between project phases or operating environments. Before transfer, organizations should reassess sensor coverage, class balance, calibration, model thresholds, user roles and intervention authority against the new context. Validation should include abnormal and degraded conditions, not only representative normal operations, because safety systems are most valuable when circumstances depart from expectation. Performance evidence should be reviewed jointly by data specialists and domain experts, with disagreements documented rather than averaged away. Where local evidence is insufficient, the model should revert to advisory status and established engineering controls should govern the decision. This preserves learning while preventing an impressive but weakly validated model from acquiring authority faster than the evidence justifies. Periodic review should confirm users understand meaning, fallback procedures remain workable, and conditions have not introduced new errors.

Practical and Managerial Implications:-

Organizations should begin with high-value safety decisions rather than with available technology. A suitable use case has a defined hazard, measurable precursor, decision window and control response. Pilot deployment should compare AI-assisted decisions with existing practice and record not only accuracy but also response time, workload and unintended effects. Model owners, data owners and risk owners should be explicitly named.

Procurement specifications should require evidence about training data, validation conditions, sensor reliability, cybersecurity, confidence reporting and maintenance. Contractor data should be governed transparently because surveillance technologies can create resistance when purpose and access are unclear. Table 2 provides an assurance checklist covering data integrity, model validity, human oversight, control linkage and lifecycle governance. AI should enhance professional capability while retaining fallback procedures that allow safe work during outages or model uncertainty.

Limitations and Future Research:-

The literature remains dominated by prototypes, controlled demonstrations and retrospective datasets. Direct evidence that AI reduces severe incidents is limited because such events are rare and field evaluation is difficult. Future research should use prospective multi-site designs, report calibration and false-negative performance, and evaluate human response to alerts. More work is needed on contractor-heavy settings, abnormal operations, extreme weather and transfer across cultural or regulatory contexts. Studies should also compare the effectiveness of human-only, AI-assisted and automated decisions rather than reporting model accuracy in isolation.

Conclusion:-

AI can strengthen proactive safety in construction and oil and gas operations when prediction is embedded in an accountable control loop. Current evidence supports computer vision, predictive analytics, language systems, sensor intelligence and digital-twin approaches, but each introduces uncertainty and new dependencies. The appropriate design principle is augmentation: use AI to improve the timeliness, coverage and traceability of risk information

while preserving engineering barriers, professional judgement and clear authority. A predictive safety system is credible only when data quality, model validity, human oversight, intervention logic and lifecycle governance are assured together.

References:-

1. 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. *Automation in Construction*, 164, 105443. <https://doi.org/10.1016/j.autcon.2024.105443>
2. Liu, J., Luo, H., & Liu, H. (2022). Deep learning-based data analytics for safety in construction. *Automation in Construction*, 140, 104302. <https://doi.org/10.1016/j.autcon.2022.104302>
3. Baduge, S. K., Thilakarathna, S., Perera, J. S., Arashpour, M., Sharafi, P., Teodosio, B., Shringi, A., & Mendis, P. (2022). Artificial intelligence and smart vision for building and construction 4.0: Machine and deep learning methods and applications. *Automation in Construction*, 141, 104440. <https://doi.org/10.1016/j.autcon.2022.104440>
4. Maali, O., Ko, C.-H., & Nguyen, P. H. D. (2024). Applications of existing and emerging construction safety technologies. *Automation in Construction*, 158, 105231. <https://doi.org/10.1016/j.autcon.2023.105231>
5. Larbi, J. A., Tang, L. C., Larbi, R. A., Abankwa, D. A., & Danquah, R. D. (2024). Developing an integrated digital delivery framework and workflow guideline for construction safety management in a project delivery system. *Safety Science*, 175, 106486. <https://doi.org/10.1016/j.ssci.2024.106486>
6. Hou, X., Li, C., & Fang, Q. (2023). Computer vision-based safety risk computing and visualization on construction sites. *Automation in Construction*, 156, 105129. <https://doi.org/10.1016/j.autcon.2023.105129>
7. Wen, S., Park, M., Tran, D. Q., Lee, S., & Park, S. (2024). Automated construction safety reporting system integrating deep learning-based real-time advanced detection and visual question answering. *Advances in Engineering Software*, 198, 103779. <https://doi.org/10.1016/j.advengsoft.2024.103779>
8. Kim, H., & Yi, J.-S. (2024). Image generation of hazardous situations in construction sites using text-to-image generative model for training deep neural networks. *Automation in Construction*, 166, 105615. <https://doi.org/10.1016/j.autcon.2024.105615>
9. Wang, X., & El-Gohary, N. (2023). Deep learning-based relation extraction and knowledge graph-based representation of construction safety requirements. *Automation in Construction*, 147, 104696. <https://doi.org/10.1016/j.autcon.2022.104696>
10. Alkaissy, M., Arashpour, M., Golafshani, E. M., Hosseini, M. R., Khanmohammadi, S., Bai, Y., & Feng, H. (2023). Enhancing construction safety: Machine learning-based classification of injury types. *Safety Science*, 162, 106102. <https://doi.org/10.1016/j.ssci.2023.106102>
11. Tözer, K. D., Gürcanlı, G. E., Çelik, T., & Akboğa Kale, Ö. (2024). Safer designs with BIM based fall hazards identification and accident prevention. *Safety Science*, 169, 106346. <https://doi.org/10.1016/j.ssci.2023.106346>
12. Sarvari, H., Edwards, D., Rillie, I., & Posillico, J. J. (2024). Building a safer future: Analysis of studies on Safety-I and Safety-II in the construction industry. *Safety Science*, 178, 106621. <https://doi.org/10.1016/j.ssci.2024.106621>
13. Pan, Y., & Zhang, L. (2023). Integrating BIM and AI for smart construction management: Current status and future directions. *Archives of Computational Methods in Engineering*, 30, 1081-1110. <https://doi.org/10.1007/s11831-022-09830-8>
14. Dobrucali, E., Demirkesen, S., Sadikoglu, E., Zhang, C., & Damci, A. (2024). Investigating the impact of emerging technologies on construction safety performance. *Engineering, Construction and Architectural Management*, 31(3), 1322-1347. <https://doi.org/10.1108/ECAM-07-2022-0668>
15. Park, J., Seong, S., Park, S., Kim, M., & Kim, H. Y. (2024). Multi-label material and human risk factors recognition model for construction site safety management. *Journal of Safety Research*, 91, 354-365. <https://doi.org/10.1016/j.jsr.2024.10.002>
16. Sanni-Anibire, M. O., Mahmoud, A. S., Hassanain, M. A., & Salami, B. A. (2020). A risk assessment approach for enhancing construction safety performance. *Safety Science*, 121, 15-29. <https://doi.org/10.1016/j.ssci.2019.08.044>
17. Koroteev, D., & Tekic, Z. (2021). Artificial intelligence in oil and gas upstream: Trends, challenges, and scenarios for the future. *Energy and AI*, 3, 100041. <https://doi.org/10.1016/j.egyai.2020.100041>
18. Sattari, F., Lefsrud, L., Kurian, D., & Macciotta, R. (2022). A theoretical framework for data-driven artificial intelligence decision making for enhancing the asset integrity management system in the oil & gas sector. *Journal of Loss Prevention in the Process Industries*, 74, 104648. <https://doi.org/10.1016/j.jlp.2021.104648>

19. Priyanka, E. B., Thangavel, S., Gao, X.-Z., & Sivakumar, N. S. (2022). Digital twin for oil pipeline risk estimation using prognostic and machine learning techniques. *Journal of Industrial Information Integration*, 26, 100272. <https://doi.org/10.1016/j.jii.2021.100272>
20. Khan, A., Al Busafi, M., Al Thuhli, A., & Siddiqui, S. (2024). Accident and incident analysis in the oil and gas sector using artificial intelligence and machine learning. *Procedia Computer Science*, 251, 295-302. <https://doi.org/10.1016/j.procs.2024.11.113>
21. Andrews, B., Chakrabarti, A., Dauphin, M., & Speck, A. (2023). Application of machine learning for calibrating gas sensors for methane emissions monitoring. *Sensors*, 23(24), 9898. <https://doi.org/10.3390/s23249898>
22. Gazos, A., Kahn, J., Kusche, I., Büscher, C., & Götz, M. (2025). Organising AI for safety: Identifying structural vulnerabilities to guide the design of AI-enhanced socio-technical systems. *Safety Science*, 184, 106731. <https://doi.org/10.1016/j.ssci.2024.106731>
23. van Dinter, R., Tekinerdogan, B., & Catal, C. (2022). Predictive maintenance using digital twins: A systematic literature review. *Information and Software Technology*, 151, 107008. <https://doi.org/10.1016/j.infsof.2022.107008>
24. Datta, S. D., Islam, M., Sobuz, M. H. R., Ahmed, S., & Kar, M. (2024). Artificial intelligence and machine learning applications in the project lifecycle of the construction industry: A comprehensive review. *Heliyon*, 10(5), e26888. <https://doi.org/10.1016/j.heliyon.2024.e26888>
25. Ajayi, A., Oyedele, L., Owolabi, H., Akinade, O., Bilal, M., Davila Delgado, J. M., & Akanbi, L. (2020). Deep learning models for health and safety risk prediction in power infrastructure projects. *Risk Analysis*, 40(10), 2019-2039. <https://doi.org/10.1111/risa.13425>
26. Kochovski, P., & Stankovski, V. (2021). Building applications for smart and safe construction with the DECENTER Fog Computing and Brokerage Platform. *Automation in Construction*, 124, 103562. <https://doi.org/10.1016/j.autcon.2021.103562>
27. Lu, H., Xi, D., & Qin, G. (2023). Environmental risk of oil pipeline accidents. *Science of the Total Environment*, 874, 162386. <https://doi.org/10.1016/j.scitotenv.2023.162386>
28. Wasim, M., & Djukic, M. B. (2022). External corrosion of oil and gas pipelines: A review of failure mechanisms and predictive preventions. *Journal of Natural Gas Science and Engineering*, 100, 104467. <https://doi.org/10.1016/j.jngse.2022.104467>
29. Shafei, H., Rahman, R. A., & Lee, Y. S. (2025). Evaluating Construction 4.0 technologies in enhancing safety and health: Case study of a national strategic plan. *Journal of Engineering, Design and Technology*, 23(4), 1211-1242. <https://doi.org/10.1108/JEDT-08-2023-0330>
30. Kim, J.-M., Yum, S.-G., Adhikari, M. D., & Bae, J. (2024). A deep-learning approach to leveraging natural hazard indicators for improved safety on construction sites. *Safety Science*, 177, 106596. <https://doi.org/10.1016/j.ssci.2024.106596>