© JUL 2026 | IRE Journals | Volume 10 Issue 1 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I1-1719466

AI-Assisted Cyber Resilience for Hazardous Gas Detection, CCTV, And Industrial Safety Monitoring Systems

SHAJJI MOHIUDDIN
ORCID: *0009-0001-8895-1614*

Abstract- ***As industrial safety systems become increasingly interconnected, the benefits of real-time visibility are accompanied by greater cybersecurity and operational risk. This paper proposes an AI-assisted Cyber Resilience model for hazardous gas detection, CCTV analytics, and industrial safety monitoring in high-risk facilities. This model is designed for high-risk industrial environments such as wellheads, petrochemical facilities, and GOSPs. It integrates multimodal safety evidence, including gas detection, video analytics, thermal imaging, and permit-control data, into a unified resilience framework, access control, digital work-permit systems, smart badges, QR-enabled certification records, air-quality monitoring (AQM)/weather stations, worker-health wearables, vehicle-permit data, and operational-technology telemetry. The paper synthesises evidence from 2020-2025 literature on real-time AI agents, industrial control security, gas-sensor analytics, computer-vision safety monitoring, and operational-technology governance. System effectiveness depends not only on model accuracy, but also on the integrity, availability, explainability, and recoverability of the full monitoring chain, but also on the integrity, availability, explainability, and recoverability of the entire monitoring chain. A resilient design therefore requires multimodal validation, edge inference, secure communications, tamper-evident evidence, human override, and staged response playbooks. The paper proposes an architecture that consolidates multimodal operational and safety signals into an AI-supported risk engine linked to OT security and safety response functions, an OT security operations function, and safety-instrumented response. It also offers a maturity model and six practical critical-area use cases covering digital work permits, working at height, confined-space entry, heavy lifting, worker safety and vehicle movement [1-8].***

Keywords: ***Artificial Intelligence, Cyber Resilience, Hazardous Gas Detection, CCTV Analytics, Digital Work Permit, Smart Badge, GOSP, Wellhead Safety, Industrial Safety, Operational Technology, Anomaly Detection, Risk Prioritisation.***

I. INTRODUCTION

Industrial facilities increasingly depend on converged safety and digital systems. Fixed gas detectors, portable personal monitors, CCTV cameras, access badges, maintenance tablets, distributed control systems, and security information platforms now exchange data across networks that were originally designed for separated operational domains. This convergence improves visibility, yet it also exposes safety functions to cyber and data-quality failures.

A forged camera stream may hide a worker entering a restricted zone; a disabled gas sensor may remove the first warning of hydrogen sulphide or methane; a ransomware event may interrupt the evidence chain needed for emergency response. Current operational-technology guidance therefore treats sensors, controllers, communications, and analytics as assets that must remain reliable under physical and cyber stress [6,7].

AI is well suited to industrial safety because hazards rarely appear as single, unambiguous signal. Gas plumes move with ventilation and weather, CCTV images degrade under glare or dust, and process values drift before equipment becomes unsafe. Machine learning can learn normal patterns, fuse heterogeneous inputs, and identify weak precursors that manual supervision may miss.

In gas sensing, multimodal fusion of sensor arrays and thermal imagery has outperformed single-modality detection in experimental settings [12,13]. In video monitoring, modern object-detection models can identify missing protective equipment, worker proximity to moving plant, and unsafe postures in real time [18-23].

In cybersecurity, anomaly models can monitor industrial protocols and process behaviour for deviations that may indicate compromise [2-5].

However, AI also introduces new dependencies. A model can be brittle outside its training domain, biased by poor labelling, misled by adversarial perturbations, or overwhelmed by alert floods.

These weaknesses are serious in safety monitoring because false negatives may permit injury while false positives can desensitise operators or trigger unnecessary shutdowns. Cyber Resilience, rather than detection performance alone, is therefore the appropriate design objective. Resilience asks whether the system can anticipate, withstand, respond to, and recover from events while maintaining acceptable safety and production states [6,8].

The topic is especially important for hazardous industrial sites such as wellhead pads, petrochemical units, gas oil separation plants (GOSPs), oil processing plants, refineries, chemical plants, compressor stations, mines, utilities, construction megaprojects, and logistics yards. Such settings combine explosive atmospheres, confined spaces, working at height, heavy lifting, vehicle movement, shift work, contractors, legacy controllers, and large visual blind spots.

This review positions AI as a decision-support capability within a governed safety architecture, not as an autonomous replacement for certified protection layers. The goal is to show how gas detection, CCTV, permit-to-work data, smart badges, AQM/weather stations, handheld monitors, worker-health sensors and vehicle authorisation can be integrated into a Cyber Resilient model that improves detection, prioritises risk, preserves evidence, and supports timely human decisions.

II. AIM AND OBJECTIVES OF THE STUDY

The aim of this review is to develop a publishable, practice-oriented model for AI-assisted Cyber Resilience in industrial safety monitoring systems. The scope covers hazardous gas detection, CCTV-based safety analytics, digital permit-to-work control, smart-badge authorisation, QR-enabled equipment and certification checks, worker-health monitoring, vehicle movement control, and connected monitoring systems used to detect unsafe conditions, verify alarms, and coordinate response.

The review deliberately focuses on the intersection of safety engineering and cybersecurity because incidents in modern facilities often cross that boundary: the same network that improves situational awareness can become an attack path, and the same AI alert that improves response can become a source of operational noise if it is not governed.

Six objectives guide the study. First, the paper analyses recent evidence on AI-enabled gas detection, video analytics, and industrial anomaly detection. Second, it identifies failure modes that weaken safety monitoring, including sensor drift, spoofing, denial of service, data poisoning, latency, poor lighting, occlusion, fragmented alarm ownership, and unverified field authorisation.

Third, it proposes a multimodal architecture that combines edge inference with secure telemetry, OT security monitoring, and safety-team workflows. Fourth, it develops a risk-prioritisation logic that ranks alerts according to confidence, consequence, asset criticality, work-permit status, personnel exposure, and cyber context.

Fifth, it formulates implementation recommendations for facilities that must maintain high availability, compliance evidence, and human accountability while adopting AI. Sixth, it converts the framework into practical use cases for wellhead areas, GOSPs, oil processing plants, confined spaces, working-at-height zones, lifting operations, worker-health monitoring and vehicle movement [1,6,7,17,24].

III. REVIEW METHODOLOGY

A structured narrative methodology was used because the topic spans safety engineering, computer vision, gas-sensor analytics, industrial control security, and governance. The search strategy prioritised peer-reviewed studies and authoritative guidance published between 2020 and 2025.

Search terms combined artificial intelligence, gas detection, sensor fusion, thermal imaging, CCTV safety monitoring, personal protective equipment detection, anomaly detection, industrial control systems, operational technology security, Cyber Resilience, and incident response.

The attached reference paper was used for structural guidance on real-time AI agents, sensor fusion, edge processing, communication reliability, safety mechanisms, and methodological organisation [1].

Sources were screened for relevance to three questions: what AI methods support industrial safety detection; what cyber or operational weaknesses can undermine those methods; and what controls enable a dependable response.

Studies were excluded when they were purely promotional, lacked methodological detail, or addressed generic enterprise monitoring without relevance to physical safety. Standards and guidance documents were included where they defined OT security requirements, governance functions, assessment domains, or risk communication practices [6-8,24].

The evidence was coded into four domains. The sensing domain captured gas sensors, thermal cameras, CCTV, access systems, digital work permits, QR codes, smart badges, wearables, AQM/weather stations, vehicle telemetry and process telemetry. The AI domain captured object detection, sensor fusion, sequence modelling, autoencoders, drift compensation, transfer learning, and explainable scoring.

The cyber domain captured segmentation, identity, secure communications, signed firmware, logging, intrusion detection, and incident response. The safety domain captured alarm verification, evacuation, shutdown, confined-space entry, working at height, critical lifting, contractor management, driver/vehicle authorisation, worker-health escalation and recovery. The synthesis then mapped each domain into a single resilience lifecycle: acquire, verify, score, act, recover, and learn.

IV. INDUSTRIAL SAFETY MONITORING AS A CYBER-PHYSICAL SYSTEM

Hazardous gas detection is a cyber-physical function because a chemical condition in the field becomes a digital signal that must be trusted before it is used for response. Industrial detectors may use electrochemical, catalytic, infrared, photoionisation, ultrasonic, or optical methods, each with different drift, cross-sensitivity, calibration, poisoning, and environmental limitations.

Machine learning can compensate for drift, improve selectivity, and classify gases from multisensor arrays, but these benefits depend on representative data and robust calibration. Recent work on gas sensor drift shows that environmental variation can be confused with instrumental degradation, which makes continuous validation essential [28-30].

CCTV safety monitoring is similarly cyber-physical. Cameras observe people, vehicles, hot work, smoke, spills, flames, barricades, and restricted areas; AI models then convert visual patterns into alerts.

Object-detection systems have demonstrated strong potential for identifying helmets, vests, goggles, safety shoes, and other protective equipment, yet they can struggle under dust, fog, low light, glare, long-distance views, small objects, and occlusion [18-23]. Safety cameras also create privacy, retention, and evidence-handling requirements.

The issue is not simply whether a model can detect a missing helmet but whether the detection can be trusted, explained, stored, acted upon, and reviewed after an incident.

Industrial control and monitoring networks add a third layer of complexity. Gas detectors and cameras may connect through gateways, industrial Ethernet, wireless networks, edge servers, cloud platforms, or security operations tools.

This creates several attack paths: stolen credentials, insecure remote access, unmanaged cameras, unsigned firmware, exposed dashboards, weak time synchronisation, and compromised vendor support channels. OT guidance emphasises that industrial

systems have unique performance, safety, and reliability requirements and cannot be secured by copying enterprise IT practices without adaptation [6].

The Cyber Resilience problem is therefore a chain problem. A strong AI model provides little benefit if the input stream is spoofed, the alert server is unavailable, the safety team cannot verify the scene, or the response action is not documented.

Conversely, a modest model can add value if it is embedded in a resilient chain with redundancy, drift monitoring, calibrated thresholds, escalation rules, and human confirmation. This review therefore treats AI outputs as decision-support evidence that must be fused with process context, asset criticality, and operational constraints.

V. AI-ASSISTED DETECTION AND VERIFICATION

AI-assisted detection works best when it is designed as a layered verification system. The first layer is sensing: fixed gas detectors, portable monitors, CCTV, thermal cameras, acoustic sensors, AQM/weather stations, smart badges, QR-coded certificates, worker-health wearables, vehicle permits, process historians, and network telemetry collect signals.

The second layer is pre-processing: timestamps are aligned, corrupted values are removed, frames are de-identified when appropriate, and sensor health, certificate validity, permit status and badge identity are checked.

The third layer is inference: CNN and YOLO-family detectors analyse images, temporal models analyse sensor sequences, and autoencoders learn normal behaviour. The fourth layer is fusion: the platform cross-checks whether a gas alarm is consistent with video evidence, ventilation status, wind direction, work permits, authorised personnel, lifting certificates, vehicle movement and process conditions [12-17].

For hazardous gases, AI should not replace certified detectors or gas-monitoring procedures. Its stronger role is to reduce uncertainty around alarms. A sudden hydrogen sulphide reading near a pump seal has different meaning if CCTV shows a maintenance team opening equipment, thermal imaging shows a plume, and process telemetry shows pressure instability. Multimodal sensor fusion has shown that combining sensor-array and thermal-image data can improve gas identification compared with individual modalities [12].

Subsea and infrared gas leak research further shows the potential of computer vision to detect gas movement where human observation is weak [14,15].

For CCTV, the main design choice is whether analytics run at the edge, in a local data centre, or in the cloud. Edge inference reduces latency and preserves operations during connectivity loss, which is vital for high-risk areas. Centralised analytics can support fleet-wide model governance and long-term retraining.

A hybrid architecture is therefore appropriate: safety-critical detections run locally, while aggregated evidence and anonymised learning signals are synchronised centrally. Real-time AI-agent literature emphasises the importance of edge processing, sensor fusion, communication reliability, and fail-safe pathways under dynamic conditions [1].

Verification is essential because industrial alerts are rarely binary. A PPE detector may confuse a white hardhat with glare; a gas sensor may drift; a camera may be dirty; a network anomaly may be maintenance activity; a person in a restricted area may be an authorised emergency responder.

The review therefore proposes confidence scoring rather than raw alarms. Confidence should reflect model probability, sensor health, environmental conditions, historical false-alarm behaviour, and corroborating evidence.

Alerts below a confidence threshold should be queued for review or paired with a request for additional evidence; alerts with high confidence and high consequence should trigger immediate escalation.

© JUL 2026 | IRE Journals | Volume 10 Issue 1 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I1-1719466

VI. CYBER RESILIENCE ARCHITECTURE

Figure 1 presents the proposed architecture. Its design principle is separation with controlled integration. Field sensors and cameras are protected in operational zones; edge devices perform low-latency inference; digital work-permit, smart-badge, AQM/weather, QR-certificate, vehicle and worker-health data provide operational context; a resilience core combines safety consequence and cyber context; and response functions coordinate human action, containment, and recovery.

Security controls wrap the architecture rather than being added at the end. These controls include strong identity, least privilege, network segmentation, encrypted telemetry, signed firmware, managed configuration, secure time sources, centralised logging, backup communication routes, tamper-evident permit records, and tested incident playbooks [6-8,24].

Figure 1. AI-assisted safety sensing and Cyber Resilience architecture for hazardous gas, CCTV, and industrial monitoring systems.

VII. CYBER RESILIENCE ARCHITECTURE CONTINUED

The architecture deliberately separates safety judgement from cyber evidence while allowing both to inform prioritisation. A gas alarm near a confined-space entry may be safety-critical even when there is no sign of cyber compromise.

A camera outage near a critical compressor may be cyber-suspicious even when no gas alarm is present. The resilience core evaluates both perspectives, reducing the risk that cybersecurity teams ignore physical consequence or that safety teams overlook data integrity.

The first control layer is asset visibility. Every detector, camera, edge device, model version, gateway, dashboard, account, integration, digital permit workflow, smart badge, handheld gas monitor, AQM/weather station, QR certificate record, lifting-equipment record and vehicle-permit record should be inventoried and mapped to safety functions.

This mirrors the governance emphasis in recent cybersecurity frameworks, which stress the need to understand assets, suppliers, risk ownership, and prioritised controls [7,8]. The second layer is trustworthy telemetry.

Streams should be authenticated, encrypted where feasible, timestamped, and monitored for missing data, unusual compression, abnormal frame rates, impossible sensor values, expired certificates or inconsistent badge-location events. The third layer is model assurance. Models should have documented

training data, performance boundaries, drift metrics, approval status, and rollback options.

The fourth control layer is response orchestration. AI alerts must be routed to the right role: control-room operator, safety officer, emergency response team, OT security analyst, maintenance supervisor, or plant manager.

Playbooks should define when to verify, evacuate, isolate a network segment, bypass a camera, send a field team, start emergency ventilation, or initiate a safe shutdown. The fifth layer is recovery and learning. After the event, the organisation should verify data integrity, restore systems, review response timing, adjust thresholds, retrain models if justified, and update procedures.

VIII. EVIDENCE-BASED MAPPING OF SAFETY AND CYBER FUNCTIONS

The reviewed literature supports the need for integrated mapping because each monitoring domain has different strengths and vulnerabilities. Gas sensors provide direct chemical evidence but can drift or fail silently. CCTV provides contextual evidence but is vulnerable to visibility problems and privacy constraints.

Process telemetry provides operational context but may be compromised through OT pathways. Cybersecurity tools provide threat evidence but may not understand safety consequence. The value of AI lies in connecting these partial views without pretending that any single view is complete.

Table 1. Mapping of AI-assisted safety functions to Cyber Resilience requirements.

Monitoring domain	Typical data	AI contribution	Resilience requirement
Hazardous gas detection	Fixed detectors, portable monitors, thermal images, weather and ventilation data	Classification, drift compensation, plume confirmation and false-alarm reduction	Calibration evidence, sensor-health monitoring, authenticated telemetry and local alarm continuity
CCTV safety analytics	Video streams, thermal cameras, access badges and work-permit context	PPE detection, zone intrusion, man-down, smoke, flame, spill and vehicle proximity alerts	Privacy controls, tamper detection, camera-health checks, edge inference and evidence retention
Process and OT telemetry	PLC values, historian data, alarms, remote access and network flows	Process anomaly detection, cyber intrusion detection and event correlation	Network segmentation, signed firmware, time synchronisation, secure logging and incident playbooks
Emergency coordination	Alarm panels, radios, evacuation status, maintenance tickets and response logs	Risk ranking, escalation routing, response tracking and post-incident learning	Role clarity, backup communications, tested overrides, recovery verification and accountable reporting
Permit-to-work, access and vehicle control	Digital work permits, smart badges, QR-coded certificates, worker competency records, vehicle permits, area classifications and route data	Authorisation validation, location matching, permit-condition checks, vehicle proximity alerts and non-compliance escalation	Identity assurance, tamper-evident permit logs, QR/app verification, offline permit access, role-based approval and command-centre visibility

© JUL 2026 | IRE Journals | Volume 10 Issue 1 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I1-1719466

IX. RISK-PRIORITISED RESPONSE MODEL

Detection alone does not create resilience. A facility becomes resilient when it can convert uncertain signals into proportional action. Figure 2 therefore frames response as a five-stage workflow: acquire, verify, score, act, and learn. This workflow aligns with modern risk frameworks in which continuous monitoring, adverse-event analysis, incident management, mitigation, recovery, and communication are treated as coordinated outcomes rather than isolated tools [7].

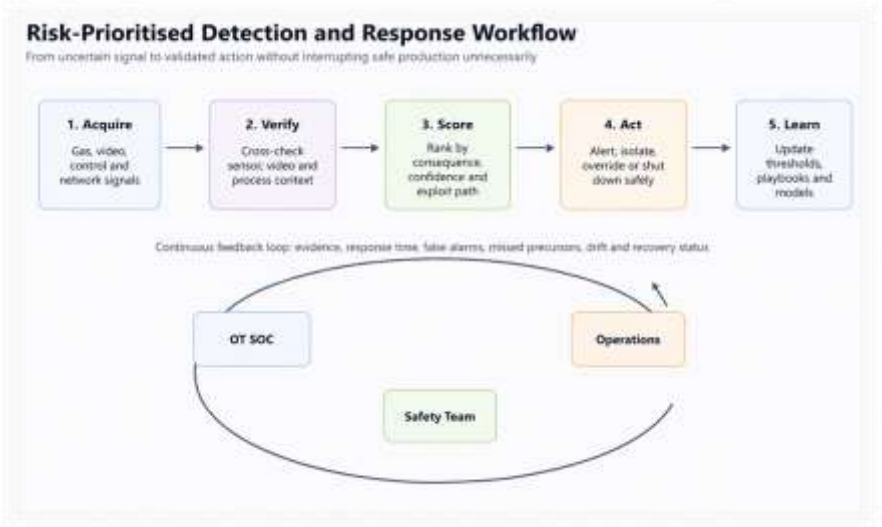

Figure 2. Risk-prioritised workflow linking AI detection, cross-domain verification, operational response, and continuous improvement.

X. RISK-PRIORITISED RESPONSE MODEL CONTINUED

Risk scoring should combine five factors. Consequence estimates the credible harm if the alert is real, including fatality potential, explosion risk, environmental impact, asset damage, and production interruption. Confidence estimates the quality of evidence, including sensor health, model certainty, corroboration, certificate validity and environmental conditions.

Exposure estimates how many people, vehicles or critical assets are in the affected zone. Authorisation estimates whether the task, person, equipment and vehicle are covered by a valid permit, certification, smart badge or approved route. Cyber context estimates whether the event could be caused or amplified by compromise, such as spoofing, malware, privilege misuse, communications loss or permit-record tampering.

A high-consequence, high-exposure event should be escalated even when confidence is incomplete; a low-consequence event with strong cyber indicators may be routed to security containment.

Response design should avoid two extremes. Fully manual response is too slow for gas release, fire, or worker-machine interaction. Fully automated response can be dangerous when data are uncertain or maliciously manipulated. The better model is graduated autonomy.

Low-risk alerts create tickets and trend analysis. Medium-risk alerts request human verification and field checks. High-risk alerts trigger alarms, access restrictions, evacuation messages, or safe process actions while simultaneously notifying accountable humans. Critical alerts should activate pre-approved fail-safe logic, but the system should still record evidence and enable human override were safe.

The response model must include communication failures. Industrial incidents often occur during

storms, shutdowns, maintenance windows, or network disruptions. Edge analytics, local alarms, redundant radio links, and offline playbooks therefore remain necessary.

Cyber Resilience also requires that safety actions not depend entirely on cloud access. If the link to the central analytics platform fails, detectors and cameras should continue to support local awareness; if a camera is compromised, gas alarms and process values should still function; if the gas network fails, CCTV and work-permit data should help identify exposure.

XI. MATURITY MODEL FOR INDUSTRIAL DEPLOYMENT

Industrial organisations vary widely in readiness. Some sites still use CCTV mainly for after-the-fact investigation and gas detectors mainly as separate alarm points. Others have integrated OT monitoring, real-time analytics, and mature emergency response.

A maturity model helps leaders plan realistic improvement. The model in Table 2 describes five levels, moving from disconnected monitoring to adaptive, auditable, and risk-informed resilience. The highest level does not mean uncontrolled automation; it means that safety and cyber teams share evidence, confidence, authority, and recovery metrics.

Table 2. Maturity model for AI-assisted Cyber Resilience in industrial safety monitoring.

Level	Capability state	Operational meaning	Priority improvement
1. Fragmented	Standalone gas alarms and CCTV with limited cyber visibility	Safety evidence is available but difficult to correlate during incidents	Build asset inventory, ownership map and minimum logging
2. Connected	Sensors, cameras and alarms are networked with basic dashboards	Operators gain visibility but cyber exposure increases	Segment networks, secure identities and define response routing
3. Intelligent	AI models support gas, video and anomaly detection	Alerts are faster but may be noisy or poorly explained	Introduce confidence scoring, drift monitoring and scenario tests
4. Risk-informed	Safety and cyber evidence are fused for prioritised response	Teams coordinate using shared playbooks and criticality ratings	Automate ticketing, evidence preservation and recovery checks
5. Adaptive	Models, controls and playbooks improve after drills and incidents	The system learns while preserving accountability and fail-safe operation	Use digital twins, red-team testing and continuous assurance metrics

XII. IMPLEMENTATION CONSIDERATIONS

Successful implementation begins with hazard analysis, not technology selection. Facilities should identify the accident scenarios where AI can genuinely reduce risk: toxic gas release near personnel, combustible gas accumulation, PPE non-compliance in energised zones, unauthorised entry, man-down events, working at height, confined-space entry, heavy lifting, vehicle-person interaction, smoke or flame appearance, camera loss in critical areas, abnormal process behaviour, and permit-to-work violations.

Each scenario needs a detection source, confirmation logic, response owner, escalation time target, and recovery criterion.

10.1 Critical-Area Use Cases for Wellheads, GOSPs and Oil Processing Plants

The field use cases identified for this review strengthen the paper because they translate the Cyber Resilience model into common high-risk tasks in

© JUL 2026 | IRE Journals | Volume 10 Issue 1 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I1-1719466

wellhead areas, petrochemical units, GOSPs and oil processing plants.

In these environments, AI should not simply detect an isolated object or gas value; it should verify whether the work is authorised, whether the area remains safe, whether the worker and equipment are certified, and whether the response can be monitored from an integrated command centre. Table 3 converts these requirements into practical controls that can be included in pilot design, procurement specifications and site acceptance testing.

Table 3. Critical-area AI use cases for wellhead, GOSP, petrochemical and oil processing facilities.

Use case	Field data and AI logic	Cyber Resilience / response requirement	Expected operational value
Digital work permit and safe-area confirmation	Digital permit-to-work record, area classification, fixed and handheld gas detectors, CCTV, ventilation status, isolation status, AQM/weather readings and task timing are fused before work starts.	Permit release depends on current evidence; gas readings, timestamps, approvals and camera evidence are logged in a tamper-evident record with local fallback during network loss.	Prevents work from starting in toxic, flammable or unauthorised zones and creates a defensible audit trail.
Working at height	Cameras, smart badges, authorised-access lists, harness/PPE detection, wind limits, AQM/weather stations and trained AI modules verify that only approved workers enter the height-risk zone.	Alerts are routed to the supervisor and command centre when a non-authorised badge, missing fall-protection item or unsafe weather condition is detected.	Reduces fall risk and improves supervision of temporary scaffolds, platforms and elevated maintenance tasks.
Confined-space entry	Gas detectors, personal monitors, CCTV/thermal views, entrant/standby/rescue-team badges, ventilation status and competency records are cross-checked continuously.	The system validates oxygen, LEL and toxic-gas limits, confirms authorised persons, monitors dwell time and escalates man-down or loss-of-signal events.	Improves entry control, early warning and rescue readiness for vessels, tanks, pits and enclosed process areas.
Heavy lifting and critical lifts	Crane certificates, lifting-equipment records, load approvals, rigging QR codes, lifting permits, exclusion-zone cameras, proximity sensors and wind/weather data are analysed before and during the lift.	QR/app scanning verifies certification, rated capacity, inspection status and permitted lift plan; unsafe wind, overload risk or exclusion-zone breach triggers immediate escalation.	Prevents uncertified equipment use, overload, poor exclusion-zone control and unsafe lifting under adverse conditions.
Worker PPE, gas exposure and health monitoring	CCTV PPE detection, handheld gas monitors, IoT wearables, smart badges and physical-health indicators such as body temperature, oxygen saturation and authorised fitness-to-work status are correlated.	Area-specific PPE rules are linked to badges and permits; early warnings are sent to the worker, supervisor and integrated command centre when exposure or health thresholds are breached.	Supports early intervention for gas exposure, heat stress, medical distress, missing PPE and worker immobility.

Vehicle safety movement	Vehicle smart badges, driver authorisation, inspection certificates, route permits, GPS/RFID, CCTV, speed/proximity sensors and audible buzzers are connected to access-control and safety systems.	The platform validates vehicle documents, approved route, speed, geofence and proximity to workers; invalid permits or unsafe approach activate buzzer alerts, gate holds or control-room escalation.	Controls restricted-area movement, reduces vehicle-person interaction risk and improves traceability of mobile equipment.

These use cases also clarify how AI should interact with certified safety practice. The system may recommend escalation, restrict access, activate local alarms or preserve evidence, but final authority remains with accountable safety roles unless a pre-

approved fail-safe action is triggered. The same design supports auditability: every serious event should retain the permit, badge, gas, video, QR-code, weather, vehicle and response record needed for investigation and recovery [6-8,24].

Data governance is a second priority. CCTV and safety data can include identifiable worker behaviour, contractor movement, and sensitive production information. Clear retention rules, access controls, audit logs, and lawful-use boundaries are necessary.

Labelling processes should include safety experts because a technically correct label may be operationally incomplete. For example, a worker without a respirator may be safe in one zone and unsafe in another, depending on permits, gas status, and task conditions. This context must be captured if AI is to support defensible decisions.

Model validation should be scenario-based. Facilities should test models under night shifts, dust, rain, steam, glare, low camera angles, partial occlusion, maintenance clothing, alarms, drills, and process start-up conditions.

Gas models should be tested against drift, humidity, temperature, cross-sensitive vapours, calibration gaps, and sensor replacement. Cyber tests should include dropped packets, replayed streams, forged timestamps, password misuse, edge-device failure, and abnormal remote access. The aim is not a single accuracy number but confidence that the monitoring chain behaves safely under degraded conditions.

Human factors must remain central. AI alerts should be understandable, prioritised, and actionable. Operators need to know why an alert was raised, what evidence supports it, what immediate action is expected, and what uncertainty remains. Alert interfaces should avoid clutter and should separate life-safety alarms from advisory trends.

Training should include simulation exercises in which safety and OT security teams jointly handle mixed events, such as a gas alarm combined with camera outage and suspicious remote login.

XIII. DISCUSSION: BALANCING SAFETY, SECURITY, AND OPERATIONAL CONTINUITY

The main contribution of AI-assisted resilience is earlier and better-contextualised warning. Traditional systems often report isolated events: gas alarm, camera alert, network alert, access violation, or process deviation.

Operators must mentally integrate these under pressure. AI can assist by correlating the events and presenting a ranked hypothesis: likely leak, likely sensor fault, likely unsafe entry, likely camera obstruction, or likely cyber manipulation. This does not remove human accountability; it improves the evidence available to accountable humans.

There are important limits. Many studies report high model performance under controlled conditions, but industrial deployment is less tidy. The warning from recent industrial intrusion-detection research is relevant: impressive results on known attacks or curated datasets may not generalise to unknown attacks and operational change [5].

The same caution applies to CCTV and gas sensing. A model trained on clean imagery may fail during sand, smoke, low light, or unusual clothing. A model trained on lab gas mixtures may fail in a plant with humidity, contaminants, vibration, or ageing sensors.

The most defensible strategy is therefore layered assurance. Certified safety instruments remain primary protection. AI provides additional detection, verification, context, and prioritisation. Cybersecurity controls protect data integrity and availability.

Operators retain authority for judgement and escalation. Management reviews metrics across safety and cyber domains: true positives, false positives, missed precursors, mean time to acknowledge, mean time to contain, sensor uptime, model drift, incident evidence completeness, and recovery time. These metrics turn AI from an experimental add-on into a managed safety capability.

XIV. FUTURE RESEARCH AGENDA

Several research gaps require attention. First, multimodal industrial datasets are limited. Gas, CCTV, process, access, and network data are often stored separately and are difficult to share because of safety, privacy, and commercial sensitivity. Future work should develop anonymised, scenario-rich benchmarks that preserve temporal relationships. Second, explainable models for safety monitoring need further development. Operators should receive concise evidence summaries, not opaque scores.

Third, adversarial resilience must be tested more rigorously, including spoofed video, poisoned training data, manipulated sensor calibration, and crafted network traffic.

Fourth, edge AI for safety systems requires stronger validation. Edge devices must operate with limited compute, heat, vibration, dust, and intermittent connectivity.

Research should examine how to maintain inference quality under resource constraints without sacrificing cyber hardening. Fifth, governance models must clarify accountability when AI assists decisions that affect evacuation, shutdown, or worker discipline. Sixth, integrated digital twins could support safer testing by simulating gas dispersion, camera views, process states, and cyber events without endangering personnel.

XV. CONCLUSION

AI-assisted Cyber Resilience should be understood not as isolated automation, but as a governed system of sensing, verification, response, recovery, and learning. By linking multimodal safety evidence with cyber-context awareness, industrial operators can make faster, more defensible, and more recoverable decisions in hazardous environments.

The review shows that AI can improve gas classification, compensate for sensor drift, detect visual safety violations, recognise anomalies in industrial control behaviour, and prioritise complex events. Yet these benefits are credible only when the monitoring chain is protected against cyber compromise, environmental degradation, model drift, and poor human-machine design.

A publishable framework for this topic should therefore combine four principles. First, multimodal evidence should be fused before serious decisions are made. Second, risk prioritisation should reflect safety consequence, confidence, exposure, authorisation status and cyber context. Third, response should be graduated, auditable, and capable of local operation during communications failure.

Fourth, continuous improvement should measure both safety outcomes and cybersecurity performance. When implemented with these principles, AI can help industrial organisations move beyond passive surveillance toward predictive, trusted, and resilient safety operations.

XVI. PRACTICAL CONTRIBUTION

This review contributes a practical synthesis for high-risk facilities that want AI benefits without weakening certified safety practice. It clarifies that the strongest use of AI is not isolated automation but evidence enrichment: linking gas, video, permit, badge, QR-certificate, vehicle, worker-health,

process and cyber signals so that operators can decide faster, justify actions and recover with confidence.

The proposed model and use-case table can be used as a checklist for feasibility studies, procurement specifications, pilot testing and audit preparation. It encourages leaders to ask whether each alert is trustworthy, whether each response owner is named, whether each degraded mode is known and whether every serious event leaf recoverable evidence.

For safety managers, the framework translates technical analytics into familiar questions about hazards, barriers and emergency readiness. For cybersecurity teams, it explains why cameras, detectors, permit systems, smart badges, edge servers and vehicle-control systems should be prioritised by physical consequence, not only by vulnerability count.

For operations leaders, it links investment decisions to measurable outcomes: lower time to acknowledge, fewer nuisance alarms, clearer evacuation evidence and shorter recovery after data loss or equipment failure.

The practical value is also methodological. Pilots should begin with one or two severe scenarios, such as toxic gas release during maintenance or missing respiratory protection inside a controlled zone, and should then test the full chain from sensing to recovery.

Procurement documents should require model cards, latency limits, offline behaviour, secure update mechanisms, integration with alarm management and evidence export. Site acceptance tests should include dirty lenses, low light, sensor replacement, communication loss, simulated spoofing and mixed human-cyber incidents.

Governance boards should review false positives, missed precursors, override decisions and privacy complaints with the same discipline used for process safety indicators. In this way, AI-assisted monitoring becomes a managed assurance programme rather than a collection of cameras and dashboards.

The essential message is clear: resilient industrial safety monitoring remains measurable, explainable, proportionate, auditable, local, secure, redundant, contextual, timely, governed, human-centred and continuously improved across sensors, models, networks, teams, suppliers, drills, incidents, recovery and learning.

Finally, the framework supports publication-quality evaluation because it separates detection accuracy from resilience performance. Future studies can report precision and recall, but they should also report latency, sensor uptime, drift rate, degraded-mode behaviour, time-to-containment, evidence integrity and operator workload.

These measures make comparisons fairer across plants with different hazards and maturity. They also prevent overclaiming by showing whether AI continues to help when conditions become noisy, adversarial or operationally constrained. That evidence is essential for regulators, insurers, executives and workers who must trust automated safety support in practice.

REFERENCES

[1] Durlik, I., Miller, T., Kostecka, E., Kozlovska, P., and Slaczka, W. (2025). Enhancing safety in autonomous maritime transportation systems with real-time AI agents. Applied Sciences, 15, 4986.

[2] Bhamare, D., Zolanvari, M., Erbad, A., Jain, R., Khan, K., and Meskin, N. (2020). Cybersecurity for industrial control systems: A survey. Computers and Security, 89, 101677.

[3] Abdelaty, M., Doriguzzi-Corin, R., and Siracusa, D. (2020). DAICS: A deep learning solution for anomaly detection in industrial control systems. Proceedings of IEEE NetSoft.

[4] Umer, M. A., Junejo, K. N., Jilani, M. T., and Mathur, A. P. (2022). Machine learning for intrusion detection in industrial control systems: Applications, challenges, and recommendations. International Journal of Critical Infrastructure Protection, 38, 100516.

[5] Kus, D., Wagner, E., Pennekamp, J., Wolsing, K., Fink, I. B., Dahlmanns, M., Wehrle, K., and

Henze, M. (2022). A false sense of security? Revisiting machine-learning-based industrial intrusion detection. Proceedings of the ACM Workshop on Cyber-Physical Systems Security.

[6] Stouffer, K., Pease, M., Tang, C., Zimmerman, T., Pillitteri, V., Lightman, S., Hahn, A., Saravia, S., Sherule, A., and Thompson, M. (2023). Guide to Operational Technology Security, NIST SP 800-82 Revision 3. National Institute of Standards and Technology.

[7] National Institute of Standards and Technology. (2024). The Cybersecurity Framework 2.0. NIST CSWP 29.

[8] National Cybersecurity Authority. (2025). Essential Cybersecurity Controls ECC-2:2024. Riyadh: National Cybersecurity Authority.

[9] Cybersecurity and Infrastructure Security Agency. (2023). Cross-Sector Cybersecurity Performance Goals. Washington, DC: CISA.

[10] Aslam, M. M., Khan, M. A., Iqbal, J., and Khan, S. (2025). An optimized anomaly detection framework in industrial control systems. Scientific Reports, 15, 12775.

[11] Seo, J. K., Lee, H., and Park, S. (2025). AI-based anomaly detection in industrial control and cyber-physical systems: A data-type-oriented review. Electronics, 15, 20.

[12] Narkhede, P., Walambe, R., Mandaokar, S., Chandel, P., Kotecha, K., and Ghinea, G. (2021). Gas detection and identification using multimodal artificial-intelligence-based sensor fusion. Sensors, 21, 3651.

[13] Sharma, A., Chandel, P., Walambe, R., Kotecha, K., and Ghinea, G. (2024). Gas detection and classification using multimodal data and federated learning. Sensors, 24, 5695.

[14] Zhu, H., Li, J., Wang, Y., and Zhang, H. (2023). Advanced computer-vision-based subsea gas leaks monitoring. Sensors, 23, 2637.

[15] Wang, M., Li, Z., Liu, Y., and Chen, X. (2025). Infrared imaging detection for hazardous gas leakage using background information and improved YOLO networks. Remote Sensing, 17, 1030.

[16] Chew, B. K., Lee, J. H., Tan, C. K., and Wong, Y. C. (2025). Autonomous hazardous gas detection systems: Sensors, analytics and safety implications. Sensors, 25, 6754.

[17] Khurram, M., Aslam, M., Ullah, I., and Khan, S. (2025). Artificial intelligence in manufacturing-industry worker safety: A new paradigm for hazard prevention and mitigation. Processes, 13, 1312.

[18] Barlybayev, A., Sharipova, A., and Omarov, B. (2024). Personal protective equipment detection using YOLOv8 for industrial safety monitoring. Cogent Engineering, 11, 2333209.

[19] Liu, H., and Qin, X. (2024). Target detection of safety protective gear using improved YOLOv5. arXiv preprint arXiv:2408.05964.

[20] Ding, Y., and Luo, X. (2023). Personal protective equipment detection in extreme construction conditions. Automation in Construction, 156, 105095.

[21] Al-Khiami, M. I., Al-Shammari, A., and Al-Busaidi, A. (2024). Development of a custom YOLOv8 model for PPE monitoring in construction environments. European Conference on Computing in Construction Proceedings.

[22] Rahman, A., Islam, M. R., and Hossain, M. S. (2025). Deep-learning-based automated inspection of personal protective equipment compliance. Computers, Materials and Continua, 84, 2311-2330.

[23] Malaikrisanachalee, S., Chotchai, S., and Wattanapong, T. (2025). ESPCN-YOLO for personal protective equipment detection under challenging construction conditions. Buildings, 15, 1609.

[24] National Institute of Standards and Technology. (2023). Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1.

[25] International Organization for Standardization. (2023). ISO/IEC 23894:2023 Information technology - Artificial intelligence - Guidance on risk management.

[26] Zheng, H., and Paiva, A. (2021). Assessing machine learning approaches to address IoT sensor drift. Proceedings of IEEE International Conference on Big Data.

[27] Dennler, N., Schütze, A., and Vergara, A. (2022). Drift in a popular metal-oxide sensor dataset reveals the need for robust validation. Sensors and Actuators B: Chemical, 372, 132620.

[28] Yang, C., Bohlin, G., and Oechtering, T. J. (2024). Environmental variation or instrumental drift? A probabilistic approach to gas sensor drift modelling and evaluation. IEEE Sensors Journal, 24, 33088-33101.

[29] Zhang, W., Hu, S., Zhang, Z., Zheng, Y., Wang, Q. J., and Lin, Z. (2024). Unsupervised attention-based multi-source domain adaptation for drift compensation in electronic-nose systems. IEEE Transactions on Instrumentation and Measurement, 73, 1-12.

[30] Yan, J., Li, X., Wang, Y., and Liu, C. (2025). AI empowers intelligent chemical sensing systems. The Innovation Information, 1, 100014.