Understanding How Industry Uses AI in HSE-Regulated Sectors

Authors: Colin Chambers[1], Freya Anderson[1], Caleb Williamson[1], Fiona McGarry[2]

Institutions: [1]Health and Safety Executive (HSE) Science Division, UK
[2] Health and Safety Executive (HSE), Engagement and Policy Division, UK

Key words: AI use case, HSE-Regulated Sectors, Assurance frameworks, Anonymous AI use case survey, Generative AI, Workforce deskilling, Computer vision.

Abstract

The Health and Safety Executive (HSE) Science Division has undertaken research to explore the uses of artificial intelligence (AI) within industries under HSE's regulatory remit. This study aimed to document real-world AI use cases, assess their potential health and safety impacts and capture assurance and control measures for safe implementation. Data collection involved HSE staff contributing use case examples as well as an anonymous industry survey, yielding 247 AI use cases across 14 sectors. It was noted that the survey respondents varied significantly in their exposure to and knowledge of AI-based systems and their assurance.

An AI classification schema was developed to categorise AI use cases based on factors including system output, data input type, implementation scale and real-time capabilities.

Artificial intelligence is used in several key areas, including maintenance systems such as advanced inspection and failure monitoring, health and safety management through risk assessments and incident analysis. AI is also used for equipment and process control, such as autonomous vehicles and robotics and for occupational monitoring using computer vision for workplace safety. Generative AI tools, such as large language models, were highlighted as being widely used to support risk assessment.

The research identified potential risks associated with AI adoption, including over-reliance on AI safety systems, workforce deskilling, inaccuracies introduced into safety assessments by AI and challenges in understanding AI decision-making processes. Assurance techniques and control measures to mitigate these risks were reported by respondents. Respondents also discussed AI implementation challenges they had faced.

Data were analysed using both manual methods and automated techniques, including zero-shot classification and topic modelling. These approaches expedited analysis and provided valuable insights into factors such as which AI technologies were prevalent in which industrial sectors, based on the data we collected.

The findings of this work indicate AI's growing influence on industrial health and safety while emphasising the necessity for robust assurance frameworks. This research has increased understanding of how AI is being used, which will help HSE to explore the risks and opportunities of AI use in industrial settings and support HSE's ability to regulate the application of AI in HSE regulated sectors. This will contribute to HSE's wider work on developing its regulatory approach to AI.

Introduction

The rapid advancement and adoption of artificial intelligence (AI) across various industrial sectors present both significant opportunities and novel challenges for workplace health and safety. The Health and Safety Executive (HSE), as Great Britain's regulator for workplace health and safety, recognises the need to proactively understand and address these developments. This study, undertaken by the HSE Science Division, forms part of a wider strategic initiative, to develop the regulatory approach to AI, to enable AI to be used safely and the potential benefits for health and safety realised.

This paper presents the findings of a research project designed to:

- Increase understanding of the use of AI applications in HSE-regulated sectors.
- Document real-world AI use cases.
- Identify potential occupational health and safety impacts associated with these use cases.
- Collate examples of current industry assurance and control methods for AI systems.

By fulfilling these objectives, this study provides a crucial evidence base to inform HSE's future policy and regulatory activities concerning the safe deployment of AI.

Methodology

This research employed a mixed-method approach to data collection and analysis to ensure a comprehensive and robust understanding of AI applications in HSE-regulated sectors.

Data Collection

Data were collected from two primary sources:

- Internal HSE staff consultation: A targeted consultation was undertaken with HSE staff from across the organisation. They were given the partially populated AI use case database and asked to add any examples they had knowledge of, completing the database fields with as much detail as they could. This internal knowledge was invaluable for identifying use cases and understanding the unique contexts of various sectors.

- Anonymous industry survey: An anonymous online survey was distributed to industry professionals via platforms such as LinkedIn. The survey invited participants to describe their use of AI systems, the associated health and safety applications and the challenges and control measures they had identified. This approach yielded 150 unique, anonymised AI use cases. A total of 294 entries were received, of which 47 were deemed 'invalid' entries, duplicates, or wish-list items and were subsequently removed to ensure data quality.

Internal HSE consultation

The AI use case database was distributed internally within HSE to relevant staff who manually added AI use cases they were aware of. They also utilised their contacts within industry organisations as well as their own knowledge of AI use in industry.

Anonymous industry survey

An anonymous online survey was primarily conducted via the LinkedIn platform, supplemented by other methods such as e-bulletins and distribution through stakeholder organisations. The survey consisted of seven questions, which were derived from a more comprehensive AI classification scheme developed by the research team specifically for this project.

Although the research team couldn't pilot the questions due to time constraints, they were discussed with the HSE's AI common interest group. This group helped ensure the questions would elicit a wide range of information about AI use cases.

The questions were open-ended, encouraging respondents to provide as much free-text information as possible about each AI use case. This approach was chosen to get detailed responses, and the survey was kept anonymous to promote free and open discussion.

While the use of LinkedIn might raise concerns about how representative the responses are, we believe professionals from all HSE-regulated sectors are active on the platform. The survey received responses from 14 of the sectors regulated by the HSE showing a broad reach across the target audience.

Online question set

AI Use Case Online Survey – UK HSE Regulated Industries

1. Does your organisation currently use AI for any health and safety purposes, increased productivity or quality improvement? (Yes/No).
2. If yes, please describe the specific AI applications you use in your company/organisation's practices (e.g., AI-powered hazard detection, risk prediction, safety training simulations, health monitoring).
3. How has AI helped improve health and safety and productivity outcomes in your organisation? (e.g., reduced accidents, improved incident reporting, improved processes/efficiencies).
4. Are there any challenges you have faced in implementing AI? (e.g., data quality, integration with existing systems).
5. What do you think the potential risks of this technological application are?
6. How do you ensure the AI systems you use are reliable and unbiased?
7. (Optional) Briefly describe any plans you have for integrating AI further into your health and safety strategy.

The anonymous industry survey was distributed August–November 2024 to external stakeholders across all HSE-regulated sectors.

AI Use Case Classification Schema

To facilitate structured analysis of the collected data, a bespoke classification scheme was developed. This schema was designed to categorise each AI use case based on multiple factors, allowing for multifaceted analysis. The classification is presented in Table 1 and formed the basis for the HSE AI use case database.

AI Classification Scheme Development

To analyse participant responses systematically, a tailored AI classification scheme was developed. This scheme was designed to ensure consistent data collection whilst helping to better understand the different AI applications identified in this study.

The classification system was constructed to gather comprehensive information across all relevant aspects of AI implementations. The scheme systematically covered AI system types, data processing, industry origin and operational challenges. By creating 13 main categories, each with detailed subcategories, this complex area was systematically addressed.

The approach recognised that participants had varying levels of technical knowledge. Those with technical backgrounds could provide detailed system specifications, whilst practitioners with less technical expertise contributed valuable experiences and key insights about practical AI use. Rather than expecting each participant to answer everything, the scheme was designed to achieve comprehensive coverage through the combination of all responses.

Iterative Refinement Process

To improve both data quality and participant engagement, the approach was continuously refined throughout data collection. Initially, pre-set response options were combined with free-text boxes. However, early analysis showed that most free-text responses were simply "don't know" or "not sure" and treated as knowledge gaps. These were then added as formal answer choices, making the survey easier to complete whilst maintaining data quality.

Technical Classification Schema

The AI classification scheme, shown in Figure 1, Figure 2and Figure 3, allowed the technical features of AI implementations to be systematically examined. This approach helped identify the specific AI types and methods used in each application, including generative AI systems and deep learning approaches.

The schema covered four main technology categories: machine learning, deep learning systems, natural language processing and generative AI applications, alongside expert systems designed for specific tasks.

Operational Characteristics Analysis

Understanding differences between data inputs and outputs was crucial for determining system operation. This analysis determined whether systems ran continuously for monitoring or operated on demand, such as responding to queries. The data characteristics section captured different input types, from structured and unstructured data to time-series information. Output capabilities varied widely, including prediction functions, autonomous control systems and content creation across formats such as text, images and audio.

AI Use Case Deployment Assessment

The deployment analysis provided insights into how widely systems were implemented and their operational maturity. Implementation scales ranged from single applications to company-wide deployments, whilst maturity levels progressed from initial pilots to optimisation stages. Integration options varied from standalone systems to fully integrated solutions, with processing capabilities covering batch operations, near real-time responses and real-time functionality.

Implementation Challenges and Future Planning

As part of the classification scheme, participants were asked to identify the primary challenges faced when implementing AI systems within their organisations. These responses were categorised to understand common barriers and obstacles across sectors and implementation scales.

Table 1: AI Use Case Classification Scheme

Classification Category	Subcategories
HSE Sector	Agriculture • Bioeconomy • Chemicals • Commercial Consumer Services • Construction • Explosives • Fairground & Theme Parks • Film/Broadcasting/Theatre • Gas & Pipelines • Logistics & Transport • Manufacturing • Mines • Offshore Energy • Onshore Oil & Gas Wells • Public Services • Quarries • Sports & Leisure • Utilities • Waste & Recycling
Application Area	Maintenance • Energy Storage & Transmission • Safety & Security • Safety Management • Equipment/Process Control • Process Optimisation • Experiment Design • Product Design Tools (GenAI)
AI Technology Type	Machine Learning • Natural Language Processing • Generative AI • Expert Systems
AI Key Applications	Computer Vision • Robots & Autonomous Systems
Basic Data Input Type	Structured Data • Unstructured Data • Time Series Data
System Output/Action Type	Prediction • Classification • Optimisation • Anomaly Detection • Decision Support • Autonomous Control • Content Generation

Implementation Scale	Single Process • Production Line • Facility-wide • Enterprise-wide
Maturity Level	Pilot/Experimental • Partially Implemented • Fully Operational • Optimising
Integration Level	Standalone System • Partially Integrated • Fully Integrated
Real-Time Capability	Batch Processing • Near Real-time • Real-time

Notes:

- Structured Data: Pre-labelled data following predetermined formats (spreadsheets, CSV files)
- Unstructured Data: Non-labelled data (emails, audio files, sensor data)
- Time Series Data: Sequential data points for trend analysis and predictive modelling
- Batch Processing: Data processed in large groups for model training
- Near Real-time: Systems with acceptable delays (seconds) such as LLMs
- Real-time: Immediate response systems for autonomous vehicle control

Table 2: Excerpt from HSE's AI Use Case in Industry Database

Category	Example 1	Example 2
Industrial AI Use Case Description	GPT-4 generates safety documents with >90% accuracy, saving time & ensuring consistency.	AI on CCTV classifies PPE/traffic violations; flags risky behaviours for supervisors.
H&S Benefits (Examples: Reduced accidents, improved incident reporting)	Reduced manual errors; consistent safety documentation.	Better ground truth, improved compliance, reduced accidents.
H&S Risks (Examples: Failure in process control, stress from algorithmic management, person detection)	Over-reliance/misinterpretation of AI outputs.	Minimal new risks; main concern is overconfidence in AI replacing supervision.
Assurance (Examples: Auditing and certification, risk management frameworks, use of standards, performance testing)	N/A	Human supervisors review AI outputs; no automation of disciplinary actions.
Control measures (Examples: Human oversight, diverse redundant decision making e.g. parallel safety system, continuous data quality monitoring)	Human oversight; AI used only for low-risk tasks.	HSEQ + site management monitor; faults flagged and repaired.
Application Area	Safety management (e.g. Risk Assessment Method Statement (RAMS), reports)	CCTV object recognition
AI Technology Type	Generative AI	Machine Learning (Computer Vision)
AI Key Applications	Decision Support	Computer Vision

Basic Data Input Type	Unstructured data (text, audio)	Time series data
System Output/Action Type	Generation (text or complete documents)	Classification (CCTV events)
Implementation Scale	N/A	Facility-wide
Maturity Level		Fully operational
Integration Level		Fully integrated
Real-Time Capability	Near real-time	Real-time (CCTV monitoring)
Plans for next 3 years	Expand AI use to inspections, permits, real-time hazard detection.	Optimise camera placement, refine detection use cases.
Respondents' involvement with AI	Exploring AI for efficiency; end-user testing.	Deployed widely; trained staff & refined risk assessments.
Challenges	GDPR/legal concerns.	Connectivity, camera placement, bandwidth, blind spots, duplicate detections.

Table 2 presents two examples representing AI applications in health and safety management and occupational monitoring: a large language model and a computer vision model. Both suggested their AI models would reduce manual errors in writing critical safety documents and improve visibility of site adherence to procedural controls. However, they acknowledged risks of data misinterpretation and over-reliance on AI systems.

Example 1 explained that human oversight ensures all AI-generated documents are "rigorously analysed by management to ensure compliance with safety standards and regulatory requirements". The model is not applied to high-risk operations to "minimise potential safety risks whilst exploring the model's benefits".

Both participants discussed implementation challenges, including legal implications around GDPR for occupational monitoring.

Example 2 identified a challenge concurrent across many sectors: "achieving reliable connectivity and bandwidth across sites. This requires operators to make bespoke trade-offs between image quality and location, increasing implementation work and system deployment time."

Development of the AI database

A database was created to store and analyse information collected both from the anonymous online survey and from direct entries by internal HSE staff from across the organisation. Its structure was aligned with the survey's question categories to streamline the data review process. Table 2 presents the database format along with sample anonymous entries from industry.

Survey respondents did not always provide details to the extent shown in the examples. However, the examples serve to illustrate concepts that are discussed in the analysis. Although the survey targeted AI practitioners in HSE-regulated industries, the researchers recognised that HSE-specific terms, such as 'assurance' and 'control measures', needed to be clearly defined in the context of AI applications to ensure consistent understanding.

'Assurance' refers to the set of processes, practices and standards put in place to ensure AI systems are developed, deployed and operated safely, securely and robustly.

'Control measures' are the technical and procedural mechanisms used to manage, guide and restrict AI system behaviour and operation, helping to prevent and mitigate risks associated with AI.

Internal staff filled in database entries independently. Some participants did not answer all research questions (for example, regarding the underpinning technologies), but researchers could use submitted information and follow-up conversations to extrapolate where necessary. For the anonymous survey, further detail could only be inferred from the initial free-text description provided by the participant.

The raw database included more than 294 entries from both internal and external contributors. Researchers identified 47 entries from the external survey that were 'invalid' entries, either submitted to voice negative opinions about AI, or revealed by their content to be AI wish lists rather than real-world use cases. These were removed from the dataset.

A research version of the database, containing 247 entries, has been preserved. HSE staff continue to add reports of industrial AI applications, making the database a growing resource for scientists, policy specialists and inspectors. The database is searchable, so inspectors can determine whether a particular application has been previously observed or reported and can better understand the technology involved.

Data Analysis

The collected data was analysed using a combination of manual and automated techniques to provide both qualitative insights and quantitative trends. The textual descriptions of the use cases were initially classified manually by the research team. This was a critical step for quality assurance and for understanding the nuances of the use cases.

Following the manual review, automated techniques were used to cross-verify and further analyse the data.

This included:

- Topic Modelling: An initial attempt at automated topic modelling using a technique called HDBSCAN was undertaken. However, this yielded overlapping clusters which made it difficult to interpret and classify.
- Zero-shot Classification: To overcome the limitations of topic modelling, a zero-shot classification model was developed. This model was trained to classify the textual descriptions based on the human-derived classifications, providing a robust method for scaling the analysis.

This dual approach ensured that the findings were both comprehensive and reliable, blending expert human judgement with the efficiency of automated analysis.

Manual analysis

An initial manual classification of approximately 10% of the survey responses was undertaken to gain an initial estimation of the scope and scale of the responses and analysis effort, as well as identifying potential data classification consistency issues. From a team review of this sample, the research team identified a need for a more robust and consistent approach to the classification. As such, an automated classification and initial analysis of the data gathered via the surveys, as described in this paper, was performed.

Automated classification and topic modelling

Within the survey, users were asked to provide a free-text summary of the uses of AI by their organisation. On its own, these free-text columns are difficult to analyse due to their unstructured nature, though similarities exist between many of the responses. Grouping them based on similarities would allow analysis of distribution of AI use cases across different sectors and technology types.

Natural Language Processing (NLP) is a type of artificial intelligence focused on extracting insight using text data and can be used to support the development of a classification scheme and classify each free-text description to build a better understanding of the distribution of use cases within the collected data.

First, Topic Modelling (Grootendorst M. 2022) was used to suggest common topics which existed within the dataset. Topic modelling is known as an unsupervised classification method. This is where there is no defined end-result to classify the summaries into; the model comes to its own conclusion on how summaries should be grouped. To begin, each summary is encoded using an encoder transformer model, which converts the unstructured free-text into an n-dimensional vector (Devlin J. et al 2019) as illustrated in Figure 1.

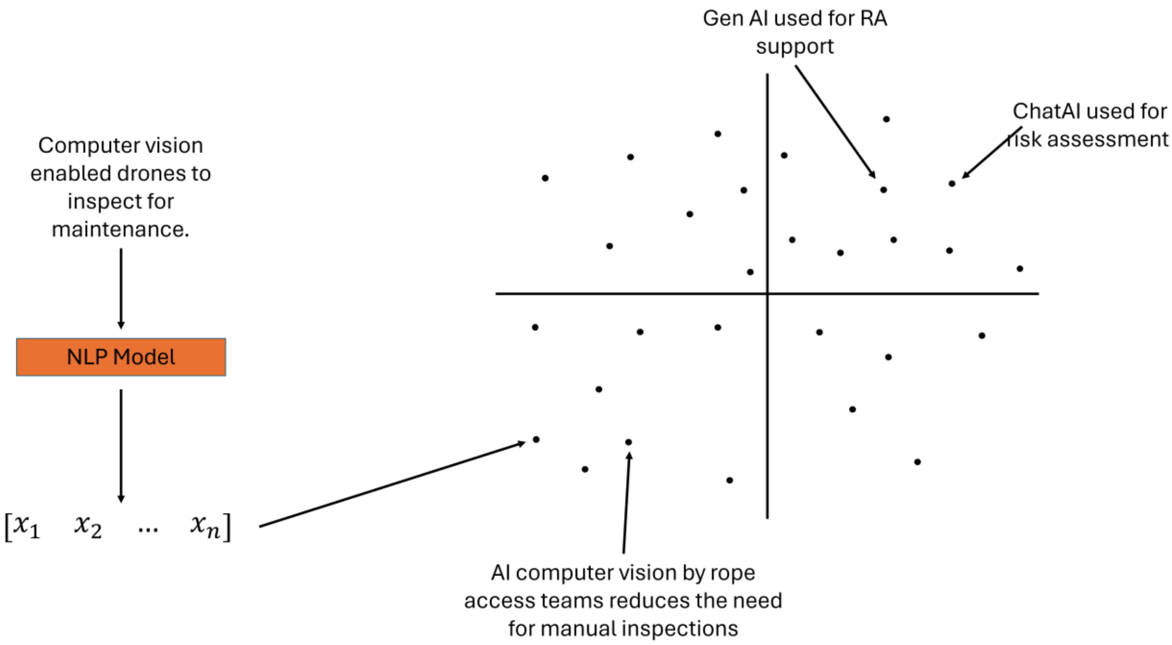

Figure 1. Embedding of free-text into an n-dimensional vector.

This approach struggled to cluster into adequate topics due to the relatively small number of use cases within the dataset. Most notably, depending on the random initialization, it would only create one cluster with the remaining documents set to outliers. It did however provide additional guidance and insight into potential classifications to use in the resulting schema. These classifications alongside a general understanding of the data were combined to produce the two classification schemes, with class descriptions, which were chosen to best represent the range of use cases existing within the data.

Next, this classification scheme was used to classify the documents using zero-shot cosine-similarity classification. Zero-shot learning is a type of machine learning where a model can classify new data without being trained on any examples of that specific category. This is a supervised classification method because the outputs are pre-defined before classification takes place.

To achieve a zero-shot cosine-similarity classification, the descriptions of each class were encoded using the same transformer model that the summaries were encoded with, such that they exist within the same vector space. The cosine similarity was then calculated between each class description and use case summary. If a summary had no class description similarity above the threshold, then it was considered an outlier and given the label 'Other'. If a summary has two or more classifications above the threshold, then the classification with the largest cosine similarity is chosen.

After evaluation of a sample of summary classifications, the threshold for classification was set to a cosine similarity score of 0.3. Figure 2 illustrates how classifications are selected based on angle from the origin between the summary and a representative class description.

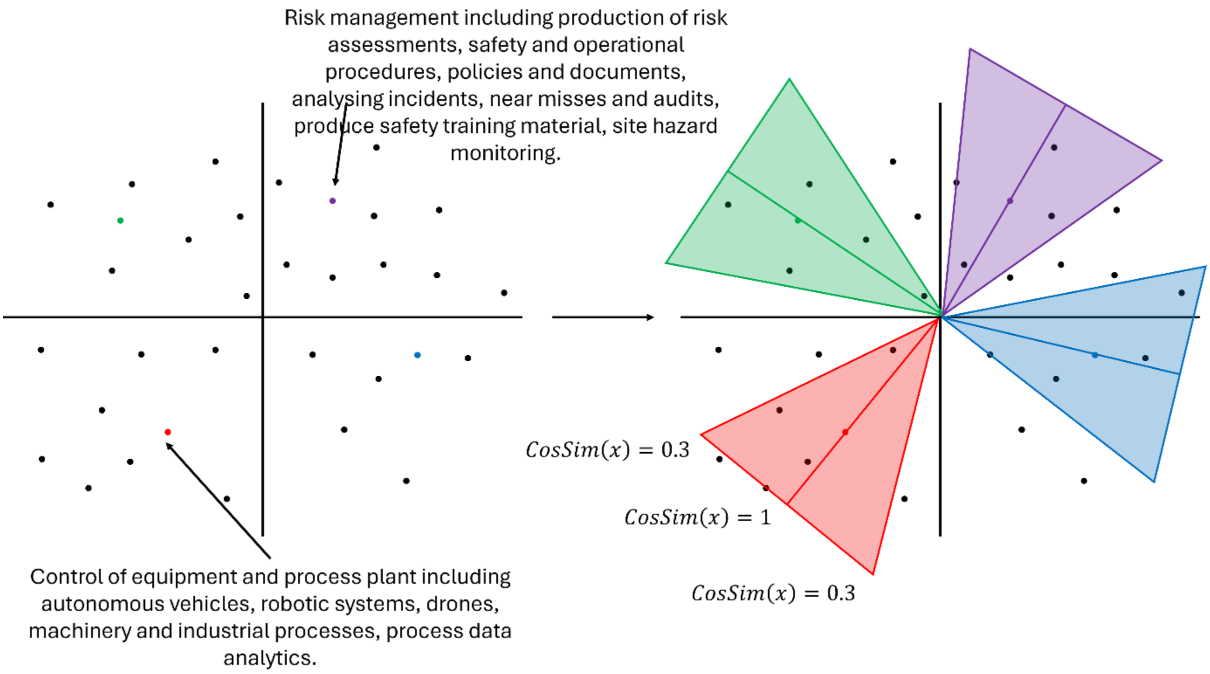

Figure 2. Zero-shot Classification Process using Cosine Similarity

The results of this classification exercise were then reviewed and the schemes refined to improve the results. This step was repeated until an adequate dataset was created.

There were two resulting classification schemes which were used for analysis, referred to as the 'high-level' classification scheme (see Table 3) and the 'low-level' classification scheme (see Table 4). The low-level classification provides a more granular view of the categories of use case provided within the data. The two schemes are as follows:

Table 3. High Level classification scheme

Title	**Description**
Maintenance Systems	Maintenance systems including monitoring for signs of equipment failure; review of inspection reports; perform automated inspection; recommend type of maintenance based on plant condition data, inspection image analysis.
Health and Safety Management	Risk management including production of risk assessments, safety and operational procedures, policies and documents, analysing incidents, near misses and audits, produce safety training material, site hazard monitoring.
Control of Plant/Equipment	Control of equipment and process plant including autonomous vehicles, robotic systems, drones, machinery and industrial processes, process data analytics.
Occupational Monitoring	Occupational monitoring, including people and object detection, PPE compliance, monitoring drivers, optimising worker behaviour, fatigue detection.

Table 4. Low Level classification scheme

Title	**Description**
Predictive Maintenance	Machine monitoring, including predictive analysis for fault and anomaly detection to support proactive maintenance.

Drone Inspections	Using AI to facilitate inspection processes using drones and advanced image analysis techniques.
Generate Health and Safety Documents	Using AI to generate and validate risk assessments, develop and optimise safety and operational procedures, policies and documents.
Analyse Incidents	Machine learning models analysing historical investigations, incident reports, near misses, audits and site observations to predict potential future events.
Generate Training Material	Generation of comprehensive training materials using AI, including videos, textual content, presentations and interactive simulations.
Control of Vehicles and Equipment	Implementation of AI and perception systems to control autonomous vehicles, robotic systems and drones.
Process Optimisation	AI-driven analysis of process parameters to ensure optimal operational efficiency.
Control of Industrial Processes	Direct AI control of industrial processes and machinery to enhance efficiency and safety.
Driver Monitoring	AI-powered monitoring of drivers and road conditions to identify operational and safety issues.
Optimise Worker Behaviour	Deployment of AI for optimising worker behaviour through management and safety monitoring systems.
Worker Health Monitoring	Integration of AI and health tracking systems to monitor fatigue.
Computer Vision	Application of computer vision for detecting objects, personnel and workers through site cameras to identify potential safety concerns.

A sample of 100 documents were selected to determine the performance of the model. For the scheme in Table 3 the model performed with an accuracy of 84 ± 7.1% and for the scheme in Table 4 it had an accuracy of 87 ± 6.6%. In most cases where the model had failed was due to summaries describing multiple use cases which caused the embedding to drift away from the corresponding correct labels and resulted in the summary becoming an outlier. The higher level likely performed worse due to the description used being much more high level and less likely to contain similar language to the more specific summaries.

Once the summaries were classified according to both schemas, the results could be compared with additional tabular data provided by the respondent. Two areas of interest to HSE were the distribution of use cases compared to sector and technology type.

Findings

The analysis of the 247 unique AI use cases revealed a diverse range of applications across 14 HSE-regulated sectors. The findings have been grouped into key areas of AI use, identified risks and the assurance and control measures currently employed.

Key Areas of AI use

Four key areas of AI application were identified. Specific examples within each category were selected to illustrate and contextualise the classification, thereby enhancing understanding and demonstrating practical relevance.

Maintenance Systems

- Drone Inspections: Drones use computer vision to inspect difficult-to-access locations and hazardous environments, such as bridges, cranes and confined spaces.
- Predictive Maintenance: AI analyses data from industrial equipment, sensors and maintenance records to recommend maintenance frequencies and actions.
- Component Failure: AI systems analyse inspection images or monitor video feeds to identify component failure signs in plant and equipment, including heavy goods vehicle inspections and fairground ride monitoring.

Health and Safety Management

- Accident Reports: Large language models analyse historical accident reports to identify hazard trends, risk factors and create incident heat maps.
- Risk Assessment: Generative AI creates risk assessments and identifies controls based on safety documentation and visual analysis, help with generating and or performing hazard and operability studies and to help with hazardous area classification calculations
- Training: AI generates health and safety induction and incident response training materials.
- Operational Documents: Large language models draft safety policies and procedures and provide live answers to health and safety queries.

Control of Equipment and Process Plant

AI is used to control autonomous vehicles, robotic systems, machinery, industrial processes and process data analytics.

- Vehicle and Equipment Control: Computer vision and sensors control equipment and autonomous vehicle movements, including quarry vehicles and farm equipment. Prevents bin lifts on refuse collection vehicles from operating when workers are in danger zones.
- Automated Operations: Autonomous systems combine robotics and algorithmic route planning. AI-driven mobile robots and forklifts manage stacking and collecting goods in automated warehouses.
- Process Optimisation: AI models analyse and optimise process plant operations, dynamically setting fill level limits and monitoring for overpressure.

Occupational Monitoring

- Safety Monitoring: Computer vision monitors worker compliance with safety procedures and protective equipment use. Monitors worker and vehicle movements, providing warnings when pedestrians and vehicles are too close.
- Workplace Monitoring: AI monitors site conditions and converts real-time video footage into searchable text for workplace activity analysis. Detects hazards like spills and leaks and alerts workers when pipelines under maintenance should not be operated.
- Worker Monitoring: AI monitors worker health data and behavioural information with process data to identify fatigue indications and exposure to vibration and sound.

Analysis against HSE original requirements

Below is an analysis of the results of the automated classification work performed.

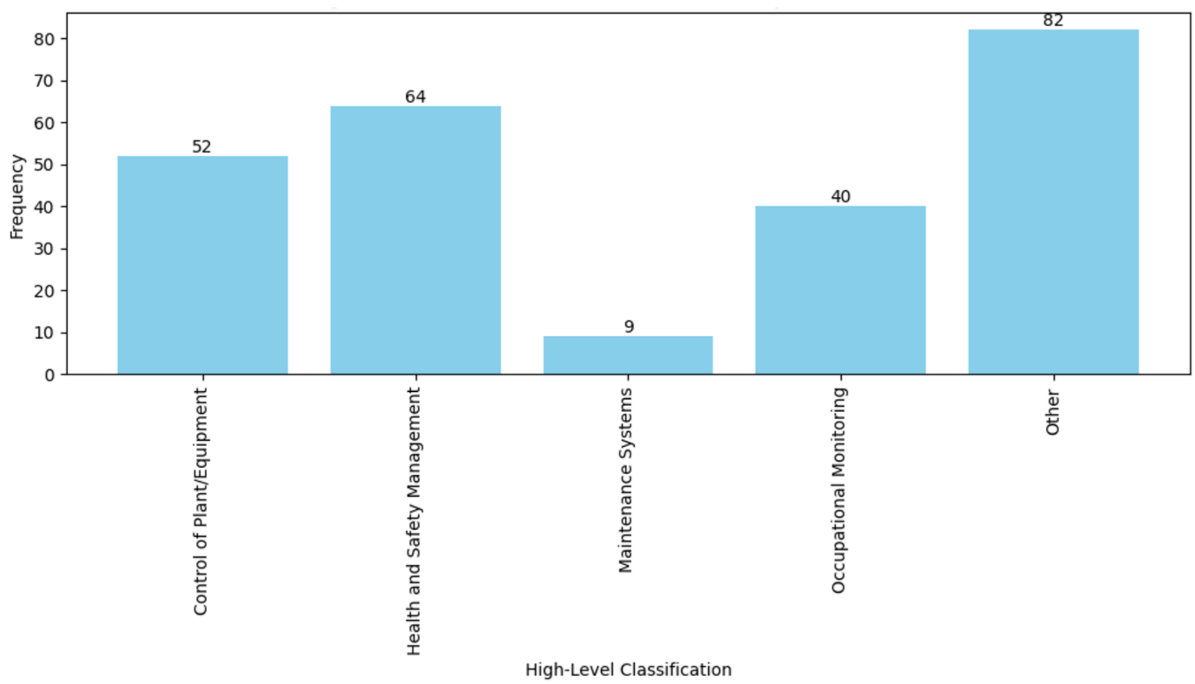

Figure 3. High-level Zero-Shot Classification summary of AI use-cases

Figure 3 shows that, based on the data we collected, more than one third of all classified examples (82 out of 247) were placed in the broad 'Other' category. Health and safety management emerged as the single largest clearly defined category at 25.9%, which could suggest that respondents saw this as an area where AI could readily deliver results. Control of plant or equipment followed closely at 21.0%, driven by the maturity of machine-learning-based control algorithms and the widespread availability of sensor data in modern machinery. Occupational monitoring, vision systems and environmental sensors accounted for 16.2%, while predictive maintenance systems remained relatively small at only 3.6%. This disparity suggests that organisations responding to the survey are still more inclined to invest in safety compliance and real-time control than in longer-term asset reliability programmes.

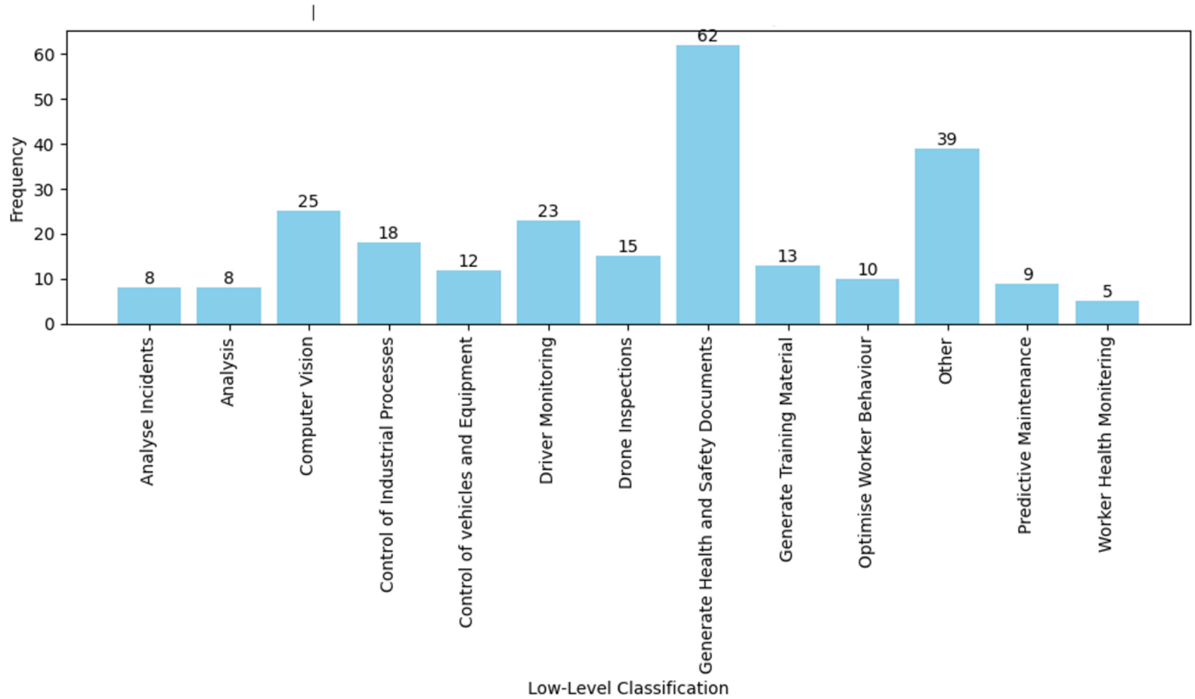

Figure 4. Low-level Zero-Shot Classification summary of AI use-cases

With reference to Figure 4, automatic generation of health and safety documents dominates the low-level landscape, capturing nearly a quarter of all examples at 25.1%. The popularity of large-language-model tools for drafting risk assessments, method statements and Control of Substances Hazardous to Health (COSHH) forms could explain this surge. Computer vision (10.1%) and driver monitoring (9.3%) together show how perceptual AI is moving from pilot to production, especially in transport and construction fleets. Drone inspections add another 6.1%, concentrated where human access is costly or dangerous, such as roofs, flare stacks and offshore decks. Incident-analysis and predictive-maintenance entries remain modest (3.2% and 3.6% respectively). Finally, the emergence of "generate training material" (5.3%) illustrates a growing appetite for personalised, AI-authored micro-learning delivered at the point of need.

Figure 5 shows that the sectors most strongly represented in the 'Control of Plant/Equipment' category are construction, logistics & transport and manufacturing, each contributing more than twice their proportional share. Occupational monitoring is similarly concentrated in these three sectors, where technologies such as wearable fatigue detectors and AI-enabled CCTV are being adopted. Logistics & transport and construction also show notable use of AI in health and safety management. In contrast, manufacturing is predominantly focused on control of plant and equipment, while traditional extractive industries (such as mining and quarrying) remain barely visible in the dataset.

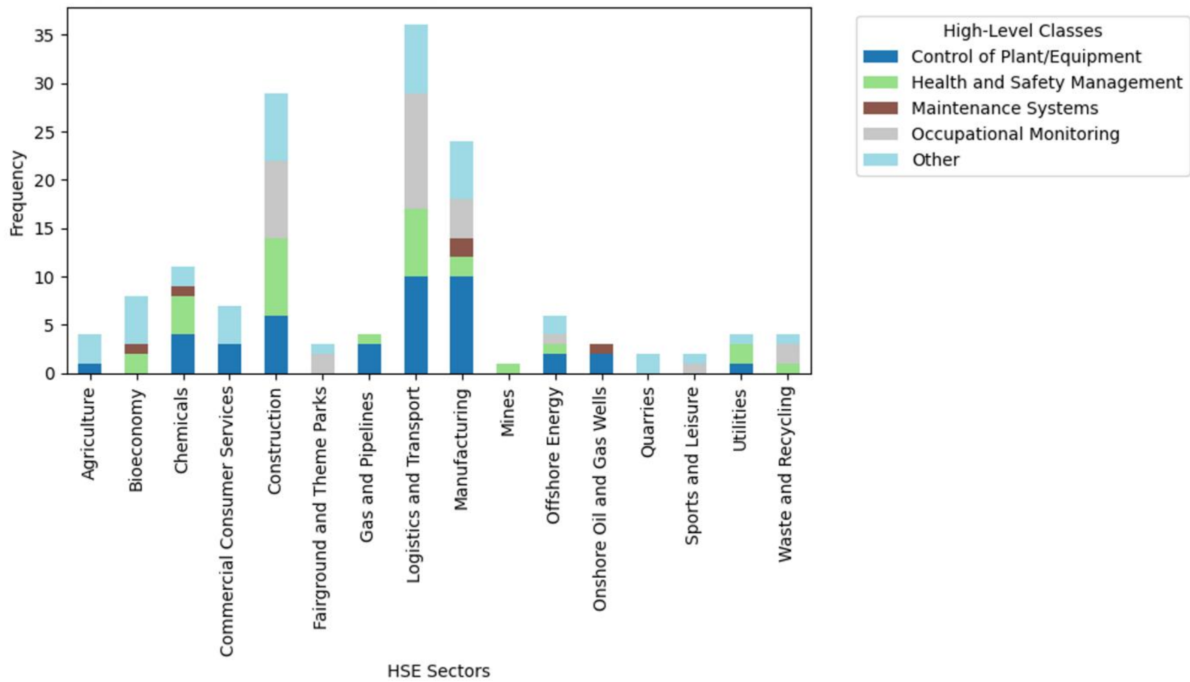

Figure 5. High-level Classification summary of AI use cases by HSE Sector

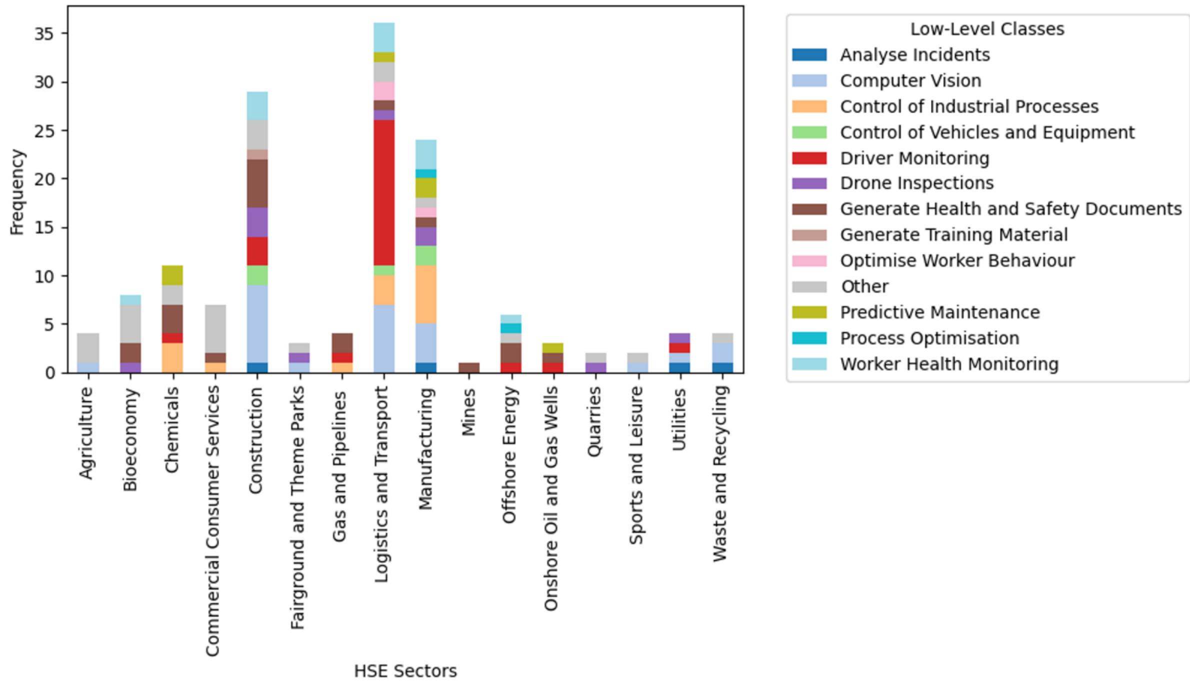

Figure 6. Low-level classification summary of AI use case by HSE sector

Figure 6 suggests that computer vision is most frequently applied in the construction sector, particularly for detecting personal protective equipment (PPE) and monitoring exclusion zones, both of which were commonly cited in the AI use cases reported to HSE. Driver monitoring is heavily concentrated in the logistics and transport sector, accounting for 78% of its total reported use. This high concentration suggests widespread adoption of fatigue-alert algorithms by long-haul fleets in those cases reported in this survey.

Drone inspections are primarily employed in settings where physical access is challenging, such as in quarries, offshore topsides and large construction projects. In contrast, document generation is more evenly distributed, with the chemicals, construction, sectors each having a significant share. This confirms the broad appeal of generative text tools across various industries with vastly different risk profiles.

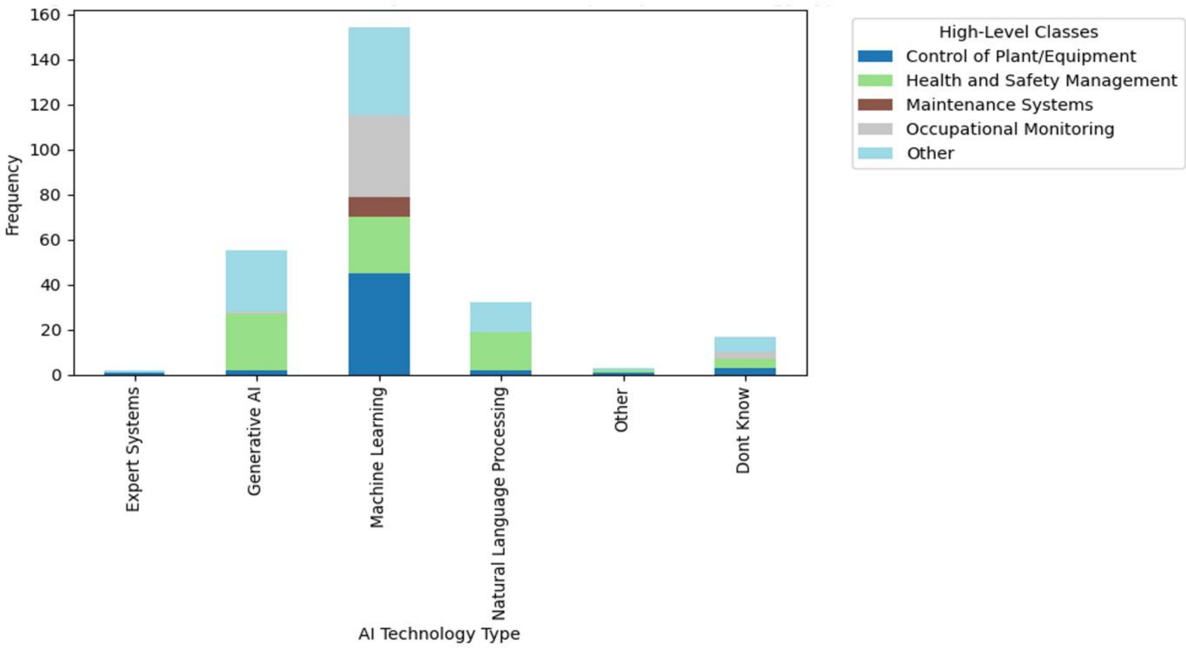

Figure 7. High-level classification summary of AI use cases by AI technology type

Figure 7 suggests that machine learning is the workhorse for every high-level theme except health and safety management, where generative AI (62.8%) and natural-language processing (42.5%) together dominate. This split makes sense: drafting policies, risk assessments and method statements are a text-heavy task perfectly suited to large-language models, whereas real-time control or monitoring still relies on numerical sensor feeds and supervised learning. Maintenance systems are almost purely ML (100%), hinting at vibration, temperature and oil-analysis datasets that lend themselves to anomaly-detection algorithms. The persistent "Don't know" slice (6–11%) could suggest that many purchasers are buying AI-enabled products without a clear understanding of the underlying technology. This could be a governance gap that auditors may query.

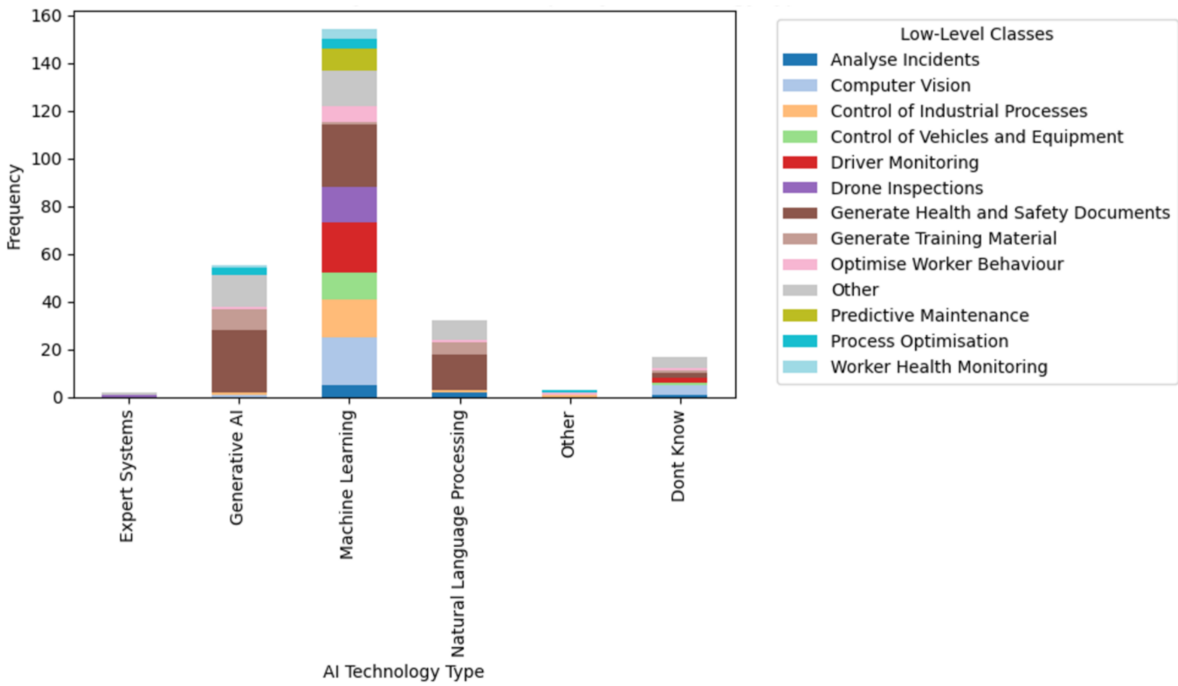

Figure 8. Low-level Classification summary Of AI use cases by AI technology type

Figure 8 shows that perceptual tasks, such as computer vision, driver monitoring and drone inspections, are dominated by machine-learning models trained on image or telemetry data (74–94%). In contrast, generative AI accounts for 61.5% of document creation and 93.3% of training-material generation, which can be blended with natural-language processing for templating or translation. The notably higher "Don't know" response for computer vision (23.5%) suggests that end-users frequently procure turnkey camera systems without realising that convolutional neural networks sit inside the box. Expert systems linger only in niche drone applications (6.3%), typically rule-based flight-log analysers. Overall, the low-level picture confirms a clear technological divide: perceptual AI is machine-learning territory, linguistic tasks use generative AI and many AI users remain uncertain about what is under the bonnet.

Identified Risks and Assurance Methods

This section presents risks and assurance controls typically cited by respondents.

Risks reported by respondents

Survey respondents identified significant health and safety risks associated with AI adoption, which were categorised by the research team into three primary areas:

Human Factors Risks:

- Over-reliance on AI systems: Potential reduction in worker vigilance, leading to complacency and inappropriate trust in AI outputs, with consequent workforce deskilling as safety-critical expertise declines.
- Psychological impacts: Continuous algorithmic monitoring may increase worker stress levels
- Alert fatigue: Excessive or frequent system warnings may desensitise operators, causing them to disregard genuinely critical alerts.

Health and Safety Risks:

- Erroneous safety assessments: AI model errors or biases could generate unsafe decisions with unpredictable outcomes, particularly where systems fail to recognise hazards or underestimate risks.
- Unsafe human-machine interactions: Poor system design or implementation may increase physical injury risks, especially in autonomous vehicle operations or robotic systems.

Technical Risks:

- Cybersecurity vulnerabilities: Integration of AI with critical infrastructure creates new attack vectors that could compromise worker safety through system breaches or malicious manipulation
- Data integrity issues: Flawed, biased, or incomplete training data may result in unreliable safety decisions, such as computer vision systems failing to detect hazards for certain demographic groups
- Lack of explainability: The 'black box' nature of many AI systems impedes understanding of decision-making processes, complicating accident investigation and system validation
- Unpredictable system behaviour: AI may respond inappropriately when encountering scenarios beyond its training parameters
- Data privacy concerns: Collection and processing of personal information, including monitoring and incident records, raises privacy and ethical considerations

Assurance and Control Measures

Industry respondents implemented various strategies to mitigate AI-related risks.

Human Oversight Controls:

- Human-in-the-loop approaches: Maintaining human supervisory roles for high-risk decisions, with mandatory human approval for safety-critical outputs.
- Competency requirements: Ensuring staff operating AI systems receive appropriate training and maintain relevant skills.
- Parallel safety systems: Implementing redundant, non-AI safety measures to provide backup protection.

Technical Assurance Measures:

- Rigorous testing and validation: Comprehensive system testing in controlled environments prior to operational deployment.
- Performance monitoring: Continuous evaluation of AI system outputs against expected parameters, with particular attention to safety-related functions.
- Data quality assurance: Implementation of data validation processes to ensure training and operational data meet quality and bias standards.
- Explainable AI techniques: Where feasible, deployment of interpretable AI models that provide insight into decision-making processes.

Governance and Risk Management:

- Clear accountability frameworks: Establishment of defined roles and responsibilities for AI system oversight and decision-making.
- Standards compliance: Adherence to relevant industry standards, such as ISO/IEC TR 5469:2024, ISO/IEC 23894:2023 and best practices for AI development and deployment.
- Regular audits: Scheduled reviews of AI system performance and safety outcomes.
- Fail-safe mechanisms: Implementation of automatic shutdown or fallback procedures for autonomous systems when errors are detected.

Data Protection and Security:

- Cybersecurity protocols: Encryption of training and operational data with robust access controls, additionally some respondents stated that AI was/could be used to detect potential cyber security threats and respond in real time.
- Privacy safeguards: Implementation of data protection measures compliant with relevant legislation.
- Independent validation: Third-party review of AI systems, particularly for safety-critical applications. These were stated as third-party companies, or sometimes a different part of the same company that was independent from the relevant project team. In some cases, they were simply stated as being a third party. These third parties performed audits, or their use was part of unnamed certification and audit activities.

The findings suggest that whilst AI presents significant opportunities for improving health and safety outcomes, successful implementation requires comprehensive risk management strategies that address technical, human and organisational factors.

Reported Implementation Challenges for AI Systems

Survey respondents identified barriers to successful AI implementation, which were categorised into three primary areas.

Technical Challenges:

- Data preprocessing: Ensuring data is clean, relevant, properly labelled, effectively stored and aligned with its intended purpose was identified as very time and labour intensive.
- Legacy system integration: Integrating AI systems with existing platforms and workflows presents compatibility issues.
- Algorithm bias: Ensuring AI algorithms are not biased can be difficult, potentially leading to incorrect results or decision making, such as inappropriate prioritising of safety risks.
- Initial setup complexity: The time and effort required for initial setup and customisation of AI to achieve desired functionality can slow development.
- Output verification: Respondents emphasised the need for content checking due to awareness that AI can produce incorrect results.

Human Factors Challenges:

- Employee distrust: Employees may initially distrust AI systems, especially if they feel technology is replacing human oversight or creating privacy issues.
- Skills shortage: A lack of trained competent people to work with AI was identified, coupled with the long learning curve to train staff to an acceptable level of competency.

Business Challenges:

- Financial constraints: Cost of installation and limitations of free software associated with AI development and use.
- Organisational expectations: Companies being either too cautious or overly optimistic, thinking AI can do more than it can do.
- Infrastructure limitations: Poor internet connectivity when using online AI tools.
- Data protection: Protecting company data during AI implementation and operation.
- Legal responsibility: Concerns were raised around GDPR and legal implications. These challenges are critical in ensuring AI systems are used responsibly and within the boundaries of data protection regulations.
- Regulatory uncertainty: As regulatory frameworks evolve in different jurisdictions the interplay will be a key question. Regulatory fragmentation and lack of global standards are and will be a problem for companies that operate globally, until this is addressed.

The most frequently cited issues included data quality and availability, which emerged as the most frequently reported challenges. Technical expertise gaps were also commonly cited, particularly regarding the shortage of skilled personnel capable of developing, deploying and maintaining AI systems. Integration difficulties with existing systems and infrastructure were highlighted as substantial barriers, especially in organisations with legacy technology frameworks.

Sectoral and Technological trends

AI use cases were identified in 14 sectors, with the highest reporting from logistics and transport, construction and manufacturing.

- Logistics and transport, construction and manufacturing accounted for the greatest share of use cases.
- Chemicals also featured prominently.

Future AI Implementation Pathways

Respondents were asked to predict their future pathway for AI implementation, specifically we asked what their plan was for the next 3 years regarding their adoptions of AI at their workplace. The responses were varied, with a significant number of respondents either not responding to this question or expressing uncertainty about how AI adoption would progress in their company. This was partly associated with the role of the respondent's function within the company: where they were users of AI but not in decision-making roles, they were unaware of their companies' plans. Other respondents were able to state what their companies' plans were. Given the low response rate to this question, any indication of future trend topic prevalence was not considered representative.

The reported future trends have been divided into four categories: health and safety, maintenance systems, control of plant and equipment and occupational monitoring.

Health and Safety Management:

- Integrate AI into real-time safety and incident monitoring, using advanced computer vision, health monitoring and sensors to detect hazards and unsafe behaviours instantly.
- Enhance AI-driven risk assessment by using more comprehensive data sources, including environmental factors and worker health metrics.
- Shift from basic data analysis to AI-driven predictive models that identify risk and prevent incidents.
- Develop AI-powered training programmes tailored to individual needs and behaviours to improve safety awareness.
- Develop interactive audit tools.

Maintenance Systems:

- Monitor operational equipment to remove workers from an area, to reduce exposure to dust, noise and vibration.

Control of Plant and Equipment:

- Fully integrated driverless machines and autonomous robotics.

Occupational Monitoring:

- AI tracking the occupancy rates of a building, in real time.
- Advanced occupational monitoring through worker exposure data.

Limitations of Self-Reported Information

The information used in this review was drawn from self-reported submissions provided by participants, who outlined the characteristics of their AI applications, including system type and related features. While these entries were subsequently examined by the research team, the assessments necessarily relied upon the free-text descriptions supplied by respondents. As such, the dataset reflects not only factual details but also elements of opinion and interpretation. It is also important to recognise that the participants' familiarity and expertise with AI are likely to have varied considerably, which may have influenced both the content and accuracy of their contributions. This introduces an inherent degree of uncertainty into the overall evidence base, which must be considered when interpreting the findings. Note that AI adoption evolves very quickly, so some self-reported information may already be outdated by the time this paper is produced.

Discussion

The findings of this study provide a valuable snapshot of the state of AI adoption within HSE-regulated sectors. Whilst the applications are diverse and growing, the pace of adoption is not uniform and a significant degree of caution appears to be in place, particularly concerning high-risk applications.

The most notable insight from the research is the widespread use of generative AI for safety documentation and risk assessment. This indicates that industry is seeking to leverage AI for knowledge management and administrative efficiency. However, this also raises new questions about the quality, accuracy and accountability for AI-generated safety content. The risk of workforce deskilling is particularly pertinent here, as reliance on AI for fundamental tasks may erode the critical thinking and expertise needed for genuine risk assessment.

The identified risks of over-dependence, lack of transparency and data quality issues are not unique to AI, but they are amplified by its implementation. The current reliance on 'human-in-the-loop' controls demonstrates a prudent approach but also highlights a potential barrier to full automation. For AI to be safely integrated into more critical, autonomous systems, both regulatory and assurance frameworks will be required that go beyond simple human oversight.

The study has its limitations, primarily stemming from the self-selecting nature of the survey participants. The data is indicative of AI use within the responding organisations, but it is not a statistically representative sample of all HSE-regulated industries. The level of detail and understanding varied significantly among respondents, which impacted the depth of the use case descriptions.

Conclusion

AI is a transformative technology. It can create and exacerbate health and safety risk but also has the potential to bring real benefits for health and safety. However, this potential can only be realised if robust assurance measures are in place and regulatory frameworks respond to technological change. The findings of this study provide a foundation for HSE's ongoing work to safeguard workers and the public as AI becomes ever more integral to industrial operations.

The study provides a snapshot of AI's current and emerging role in HSE-regulated industries, highlighting both its transformative potential and the need for robust assurance. Continued engagement, knowledge sharing and development of the regulatory approach, will be essential to ensure that AI adoption enhances, rather than undermines, workplace health and safety.

It is noted that AI assurance is being developed internationally (ISO, OECD, EU AI Act, etc), and HSE’s work may benefit from alignment with this work.

This work also concludes that findings are based on self-reported examples and give an indicative snapshot, not a full picture of all AI use cases across all HSE regulated sectors.

Disclaimer

This publication and the work it describes was funded by the Health and Safety Executive (HSE). Its contents, including any opinions and/or conclusions expressed, are those of the authors alone and do not necessarily reflect HSE policy.

Acknowledgements

The authors extend their sincere gratitude to all who contributed to this work. We are especially thankful to the HSE staff, whose unwavering support proved invaluable throughout. Special appreciation goes to the members of the HSE AI CIG for their dedicated engagement. We also gratefully acknowledge the time, insight and effort of Fiona McGarry (HSE EPD) and Tracy Read (HSE Science Division Business Partner), whose contributions were instrumental and much appreciated.

References

Devlin, J., Chang, M., Lee, K. and Toutanova K. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers), pages 4171–4186, Minneapolis, Minnesota. Association for Computational Linguistics.

Grootendorst, M. 2022. BERTopic: Neural Topic Modelling with a Class-Based TF-IDF Procedure. arXiv:2203.05794.

ISO/IEC., 2023, ISO/IEC 23894:2023. Information technology — Artificial intelligence — Risk management. International Organization for Standardization.

ISO/IEC., 2024, ISO/IEC TR 5469:2024. Artificial intelligence — Functional safety and AI systems. International Organization for Standardization.