FOOD FOR THOUGHT

AI for Industrial Safety

A Comprehensive Mapping of the Opportunities and Risks of Deployment

René Amalberti

n° 2026-05

TOPIC

Digital transition

Les Cahiers de la sécurité industrielle

FONCSI
Fondation pour une culture de sécurité industrielle

The *Foundation for an Industrial Safety Culture* (FonCSI) is a French public-interest research foundation created in 2005. It aims to:

- ▷ undertake and fund research activities that contribute to improving safety in hazardous organizations (industrial firms of all sizes, in all industrial sectors);
- ▷ work towards better mutual understanding between high-risk industries and civil society, aiming for a durable compromise and an open debate that covers all the dimensions of risk;
- ▷ foster the acculturation of all stakeholders to the questions, tradeoffs and problems related to risk and safety.

In order to attain these objectives, the FonCSI works to bring together researchers from different scientific disciplines with other stakeholders concerned by industrial safety and the management of technological risk: companies, local government, trade unions, NGOs. We also attempt to build bridges between disciplines and to promote interaction and cross-pollination between engineering, sciences and the humanities.

Foundation for an Industrial Safety Culture

A public-interest research foundation

www.FonCSI.org

6 allée Émile Monso – CS 22760
31077 Toulouse cedex 4
France

Email: contact@FonCSI.org

Abstract

Title	AI for industrial safety: A comprehensive mapping of the opportunities and risks of deployment
Keywords	artificial intelligence, industrial safety, opportunities, risks
Author	René Amalberti
Publication date	September 2026

Artificial Intelligence (AI) is rapidly making its way into all areas of corporate life, including the field of safety. Its development is generating high expectations, concerns, and questions regarding its deployment, governance, and organizational impact.

This document presents an overview of the opportunities and concerns associated with AI deployment through a graphical representation of six application domains, depicted as the petals of a flower: anticipating risks; distancing operators from the most hazardous situations; protecting exposed operators; enhancing human knowledge and decision-making; automating the monitoring and maintenance of technical systems; and, finally, enriching experience feedback. As part of the FonCSI initiative *Safety of the Future: "Living with" Uncertainty, Complexity, and New Expectations*, this document draws on the academic and technical literature as well as on the work of a foresight workshop that has brought together, since 2025, around twenty experts from industry and regulatory authorities, along with external contributors.

This document is an English translation of a document published in French in the same collection in June 2026, titled *La marguerite de l'IA en sécurité industrielle: Une cartographie détaillée des promesses et des risques du déploiement de l'intelligence artificielle* (Cahier de la sécurité industrielle n° 2026-02).

About the authors

This document was written by René Amalberti, director of the FonCSI and author of numerous books and articles on safety management in industry and healthcare. He is also a member of the French Academy of Technologies.

To cite this document

Amalberti, R. (2026), *AI for Industrial Safety: A Comprehensive Mapping of the Opportunities and Risks of Deployment*. Number 2026-05 of the *Cahiers de la Sécurité Industrielle*, Foundation for an Industrial Safety Culture, Toulouse, France (ISSN 2100-3874). DOI: 10.57071/aif478. Available from FonCSI.org/en.

Contents

Introduction

Context

Artificial intelligence (AI) is making rapid inroads into every area of business life, including the field of safety. Its adoption is generating both high expectations and concerns, as well as questions about its deployment, governance and organisational impact.

This document presents an overview of the promises and concerns associated with the deployment of AI, mapping them across **six application areas**, which are represented graphically in the form of a flower's petals (see page 5).

Given the rate of innovation in this area, it is clear that this representation of the major application areas of AI in industrial safety will become outdated at the petal-by-petal level of detail, but we anticipate that the macro level structure is likely to remain valid for some years. Readers of this document should bear in mind that these applications are, by their very nature, dated to mid-2026 and are all **evolving rapidly**, like any development in the field of AI.

The FonCSI "Living with" industrial expert group

This document is part of the FonCSI initiative titled *Safety of the future: "living with" uncertainty, complexity and new expectations.* It is a comprehensive and systemic reflection on the impact for safety and safety management of the major changes (climate change, digitalisation, economic and geopolitical transformations) taking place in our society.

The "living with" concept refers to a change in posture: work on safety is no longer simply a matter of anticipating and controlling identified risks, but of finding sustainable ways of coping with unstable, uncertain and constantly evolving environments. This perspective invites us to question established safety frameworks, from operational practices through to management and regulatory approaches, while incorporating new social and organisational expectations.

This document builds on the reflections initiated in an earlier FonCSI Cahier, *Safety in the era of "living with": Uncertainty, complexity and new expectations* **[Bieder et al. 2024]**, which provides the conceptual foundation for this approach. It draws both on a review of the academic and technical literature on this topic and on the contributions of a collective reflection process structured around prospective workshops.

Established in 2025, the FonCSI "Living with" industrial expert group brings together around twenty experts from companies and representatives of the supervisory authorities responsible for industrial safety, all of them patrons of the Foundation, who meet twice a year to discuss and deepen their thinking about forthcoming changes in industrial safety. The group also includes a number of external experts who are not patrons of the Foundation, including former industrial directors, the current director of the IMdR (a French non-profit organization that organizes information sharing on risk management practices), experts from the French national cybersecurity agency Anssi, and trade union representatives. Its deliberately stable composition allows the discussions to unfold over a long time frame, conducive to in-depth reflection.

This document draws more specifically on the work and discussions of the group's second prospective workshop, held in Paris in April 2026 on the subject: *The uses of AI in companies: what positive and negative safety implications are associated with them?* This meeting was organised and facilitated by the FonCSI team: René Amalberti, Corinne Bieder, Fiona Fadat, Clotilde Gagey, Caroline Kamaté, Hervé Laroche, Éric Marsden, Thomas Merlet and Jean Pariès. It brought together representatives of a range of organisations, as well as representatives of French trade unions (CGT, CFE-CGC). It follows on from discussions initiated during a first workshop, held in Toulouse in March 2025, devoted to the impacts of major transformations on industrial safety.

The following people participated in the foresight workshop in April 2026:

COSTE	Didier	Airbus
REUZEAU	Florence	Airbus
ROUGÉ	Sophie	Airbus
DEHARVENGT	Stéphane	Anssi
LARGIER	Alexandre	ASNR
MATHIEU	Pascal	CFE-CGC
MONFORT	Bertrand	CGT / CEA Saclay
NICOLAS	Guilhem	DGAC-DSAC
LABARTHE	Jean-Paul	EDF
LAUGIER	Cécile	EDF
MAGNE	Laurent	EDF
BAUMANN	Pierre	Engie
DEYDIER	Marie-Véronique	Engie
PARIS	Nicolas	EPSF
BOUCAULT	Julien	EPSF
BOISSIÈRES	Ivan	Icsi
LACOSTE	André-Claude	Icsi-FonCSI
ROUSSEAU	Luc	Icsi-FonCSI
GAY	Didier	Ineris
DECAUX	Anne-Sophie	Natran
PÉRINET	Romuald	Natran
CATARINO	David	OPPBTP
MUSCHE	Emmanuel	OPPBTP
HABERSTICH	Philippe	RTE
FRANÇOIS	Yohanna	SNCF
MACHADO-VERHEYE	Soizic	Suez
EURYALE	Anne-Laure	Systra
KLEIN	Michel	TotalEnergies
PAPILLAULT	Virginie	UIC
REPUSSARD	Jacques	IMdR
AUVRÈLE	Patrick	Ex-SNCF
DESCAZEAUX	Michel	Ex-Engie

In parallel, FonCSI draws internationally on academic experts close to the Foundation — primarily Norwegian, American, Australian and Argentinean experts — through repeated contacts and seminars held on the "Living with" theme. The results are regularly published in books with Springer in FonCSI's English-language collection «SpringerBriefs in Safety Management» and in the form of «Cahiers de la sécurité industrielle» (in English and French), all of these documents being *open access*.

Objectives of the document

This document aims to provide a **mapping of the promises** and concerns associated with the deployment of AI in industrial safety, and to identify its impacts, both positive and negative, as envisaged or already observed in the industrial sector.

Structure of the document

Chapter 1 presents the general framework proposed for understanding the promises of AI in industrial safety, structured around six "petals" of a flower, representing the six major application areas. This representation aims to provide an **overview**, both concise and structuring, of the principal areas in which AI can be expected to transform safety practice.

Chapter 2 provides a detailed reading of this flower, examining, for each area, applications that already exist or are likely to emerge in the near future. This analysis describes the **opportunities** and the reservations associated with their applications.

Finally, chapter 3 presents the **macro-level results of the prospective workshop**. It compares the benefits and drawbacks of each petal, both within each area and across the different areas. It thus aims to inform the strategic choices of industrial stakeholders, providing guidance for the **prioritisation of AI uses** in the field of safety. It also presents scenarios for the impact of its deployment on safety, based on two development curves for established technologies.

1

The framework developed by FonCSI: the “flower” of AI applications

The original framework proposed in this chapter (“the six-petal flower”) is the result of FonCSI’s analysis of current trends, and on the academic and technical literature on the topic.

The description of the details of the application areas and use cases in each petal of the flower is also work undertaken by FonCSI. The illustrative examples, in turn, come from two sources:

- The output of the second prospective workshop of FonCSI’s “Living with” industrial expert group held in April 2026 (a video of the event is available on foncsi.org).
- The academic and technical literature published in these fields, used to supplement the examples and enrich them with the main key publications.

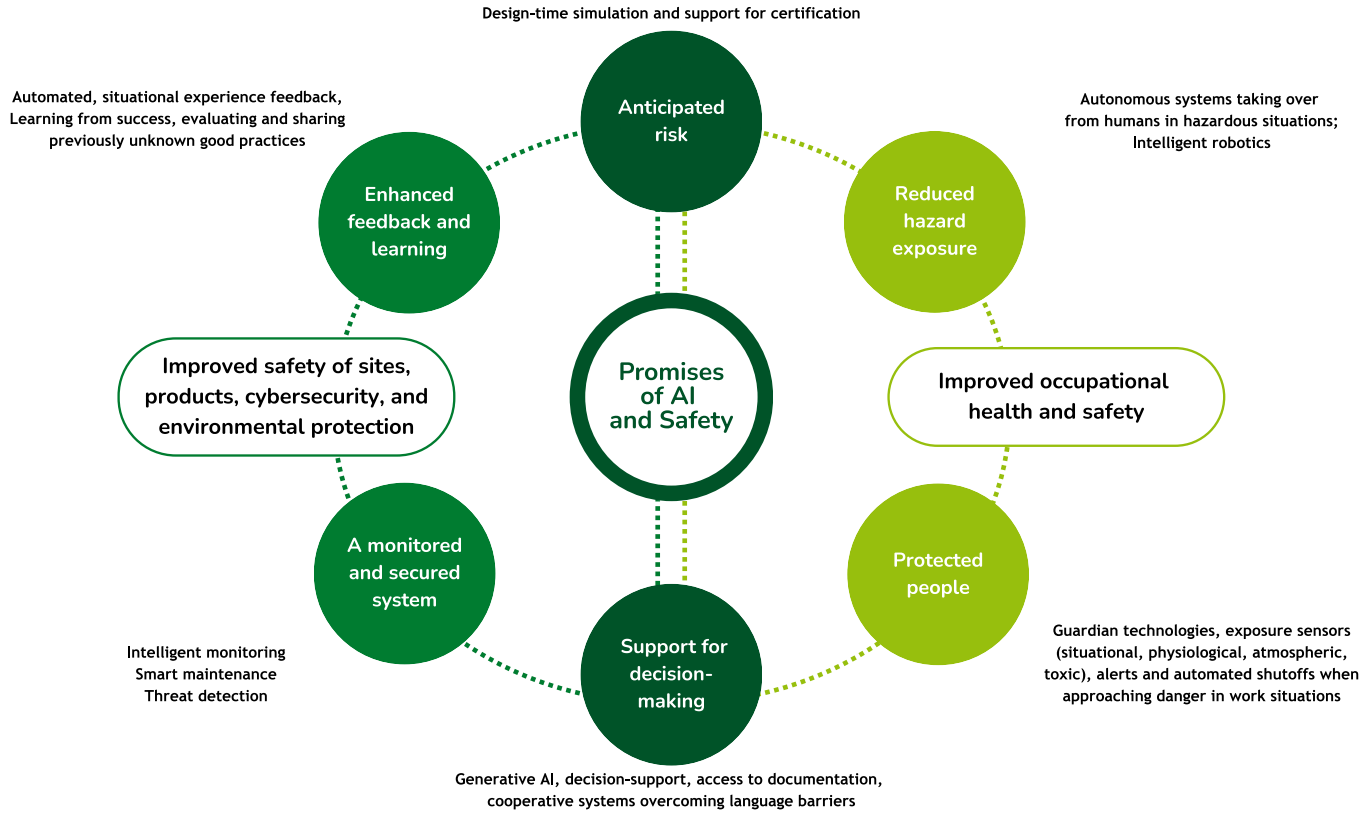

Figure 1.1 *The “flower” representing the promises for industrial safety expected to arise from the deployment of AI technologies.*

The analysis of the literature conducted prior to the group work suggests that the deployment of AI in industrial safety can be represented by **a six-petal flower** (*cf.* figure 1.1) covering the following areas:

- **anticipated risk** through safe design;
- technologies that make it possible to **reduce operator exposure to hazards**;
- technologies that make it possible to **monitor and protect the operator** when they remain directly exposed;
- technologies that **augment humans**, allowing them to be **better informed** in their everyday practices and decision-making;
- technologies that **automate** the **monitoring and maintenance** of the technical system;
- and technologies that **improve operational experience feedback**, making better use of lessons learned from daily operations and from incidents.

Although some uses of AI for safety are present in the literature, their effective deployment is not yet widespread across all industries and organizations. Furthermore, many of these uses are still at **an exploratory stage**. The "flower" therefore describes a field of possibilities as envisioned today.

2

A zoom on each flower petal

This chapter summarises the main directions for the deployment of AI and its impacts on industrial safety, petal by petal.

2.1 The 'anticipated risk' petal: AI for safe design

This area encompasses a range of application objectives.

Preliminary risk analysis and scenario exploration

AI can be used to explore broader risk-scenario spaces than can be achieved through human analysis alone. AI provides greater completeness in identifying combinations of failures, detecting 'non-intuitive' scenarios and anticipating emerging effects. Some industrial applications:

- Process industries/energy: use of AI models coupled with digital twins to **identify dangerous combinations** of parameters (temperature, pressure, ageing, human error) even before detailed design.
- Aeronautics and defence: early-stage design of decision-support or partial-autonomy functions, with **systematic exploration of 'extreme but plausible' conditions**, beyond the scenarios nominally considered by engineers.

Design verification and detection of inconsistencies

In this case, AI is used as an intelligent review tool for design models, safety requirements and system architectures, restricted to the early engineering and internal verification phases. AI is already used as a decision-support tool, but **never as the final decision-maker**. Some industrial applications:

- Aeronautics (Europe, USA): **automated consistency analysis** between safety requirements, functions and system allocations (complementary to ARP4754A / ED-79 methods), ICAO.
- Railway systems and complex industrial systems: **tools for detecting logical inconsistencies** in architecture or functional-safety models.

Inspection, validation and lessons learned (REX)

In this case, AI is used extensively on demand. In almost all cases, AI is not the safety barrier, but an additional monitoring or alerting system. Some industrial applications:

- Heavy industry, energy, infrastructure: automated visual inspection (cracks, corrosion, critical bolting) using computer vision, often via drones or robots.
- Automated plants: real-time detection of hazardous situations (human intrusion, missing PPE, risky behaviour), without being directly integrated into the certified safety function.

AI and certification

This is a more challenging area, where a distinction must be made between established uses of AI to assist certification and the major unresolved issue of certifying systems that incorporate AI technologies.

AI is already used by industry, and sometimes by authorities, to analyse dossiers, tests, lessons learned and incident histories. However, the certification of systems incorporating learning-based AI comes into direct tension with the traditional principles of certification. **Processes, safeguards and architectures are certified far more readily than 'the algorithm itself'**. A number of initiatives are nevertheless under way:

- ISO/IEC TR 5469 (2024): first attempt to establish a framework at the intersection of AI and functional safety;
- UL 4600, ISO 21448 (SOTIF) to qualify autonomous systems (robotic vehicles);
- EUROCAE WG-114 / SAE G-34 work, applied to certification in aeronautics.

2.2 The 'reduced hazard exposure' petal

This petal encompasses numerous applications of autonomous robotics whose principal safety benefit is the elimination of fatal scenarios arising from acute exposure (i.e. the elimination of human exposure to hazards). The effects are very well documented by the ILO **[ILO 2025]** and the US *National Safety Council* **[NSC 2023]**.

- In the **nuclear sector** (decommissioning, inspection, post-accident management): acute irradiation, internal contamination and unknown confined atmospheres.
  - Semi-autonomous remotely operated robots for the inspection, cutting and handling of highly radioactive waste (Sellafield, Fukushima, Chernobyl) **[Lopez et al. 2025]**.
  - Drones and quadruped robots for radiological mapping (Boston Dynamics/Spot).
  - Coupling AI with digital twins for safe planning of operations.
- In **heavy industry, chemicals and oil & gas** (ATEX zones, toxic environments): risks of **explosions, poisoning, fires and hot work**.
  - Mobile robots and remotely operated arms for the inspection, sampling and handling of hazardous substances.
  - AI-enabled drones for leak detection (gas, toxic vapours).
  - AI for continuous monitoring of critical parameters (pressure, temperature, gas).
- In **mines, quarries and tunnels**: risks of collapse, explosions and asphyxiation (oxygen-deficient atmospheres).
  - Autonomous robots for post-blast reconnaissance and tunnel inspection.
  - Teleoperation for handling explosives or debris removal.
  - Keeping operators away from the most lethal phases (post-blast, initial reconnaissance).
- In **construction and heavy handling** (AGVs, AMRs, construction robots): risks of **crushing, collision or falls**.
  - Autonomous mobile robots for transporting loads.
  - Inspection robots for work at height or in unstable areas.

The **empirically measured effects** are that a moderate increase in robot density is associated with a reduction of approximately 1.2 injuries per 100 workers per year **[Gihleb et al. 2022]**. The positive effect mainly concerns serious injuries, with an indirect impact on mortality.

2.3 The petal of the 'person protected' from proximity risks

This petal, concerning the guardian values of safety and the reduction in the severity of occupational accidents, has already been cited for several years as one of the most immediate and urgent benefits of deploying AI in the workplace (see **[Malenfer et al. 2022]**). It encompasses early warning (before human perception), the protection of lone workers, reduced reaction time, assistance with prevention rather than punishment, continuous monitoring without fatigue, and the detection of near misses.

- ▷ **'Smart' helmets or vests that perceive danger before humans**: the operator wears a helmet or vest equipped with sensors (gas, heat, noise, movements and/or location), whose signals are continuously analysed by AI. For example, if the vest detects a slight abnormal increase in toxic gas before the odour is perceptible, the system vibrates, emits an audible alert, and sends an alert to the security station. A literature review is available on applications in the construction sector and at hazardous sites **[Sakshi et al. 2024]**.
- ▷ **Proximity sensors between humans and mobile machines**: the wearing of beacons by operators and machines (forklift trucks, robots) enables AI to calculate trajectories and relative speeds.
- ▷ **Intelligent cameras which, with the help of AI, recognise dangerous situations**: an operator enters a hazardous area without a helmet; the camera detects them and issues an immediate alert. If a pedestrian approaches a forklift truck too closely, the camera detects them and alerts both the driver and the pedestrian, or likewise when a repeated dangerous action is detected (posture or fatigue).
- ▷ **Detection of fatigue, falls or discomfort**: a device combining movement, posture and physiological-parameter sensors, together with AI, can detect a fall or abnormal immobility. It also identifies signs of heat fatigue or overexertion and automatically calls the emergency services if the operator does not respond.
- ▷ **Anticipating accidents through sensors combined with AI**: AI learns from thousands of weak signals (vibrations, gestures and/or minor incidents) to provide warnings of imminent risks[1].

2.4 The 'support for decision-making' petal

The use of generative AI in the field of industrial safety is now one of the most widespread uses of artificial intelligence in safety. The number of commercially available tools is growing rapidly, and the applications are often easy to adopt (at least for superficial use) for operators who are already familiar with generative AI in everyday life[2]. There is, however, a growing body of literature concerning the concerns associated with this rapid, wide-ranging deployment, which is accompanied by little support and little testing **[Huber et al. 2025]**.

- ▷ **Situational awareness and understanding of context** (sensemaking): generative AI transforms complex flows (sensors, alarms, historical data) into simple explanatory narratives that are understandable to the operator. Applications already exist in several sectors:
  - In the nuclear and energy sectors, to provide natural-language explanations of an 'off-normal' situation (sequence of alarms, gradual drift, comparison with past incidents).
  - In chemical processes, there are narrative accounts of slow changes in parameters (temperature, pressure) that conventional thresholds have not yet detected.
  - In air and rail transport, systems are being introduced to produce an intelligible summary of several weak signals before an incident.
- ▷ **Intelligent, targeted and simplified access to technical documentation**. The objective is both to save time and to reduce procedural errors and informal workarounds. **AI becomes a cognitive intermediary** between the operator and large bodies of information (procedures, manuals, lessons learned). There are numerous examples in industrial maintenance, where AI provides rapid access to the procedures relevant to the context. Other

[1] Compliance quest, *How AI is Transforming Safety Incident Prediction*, August 28th, 2025.

[2] See, for example, the International AI safety report, 2026.

applications concern industries with stringent regulatory compliance (nuclear, oil & gas, healthcare), by providing **instant access** to the relevant parts of normative requirements, or, in older installations, by proposing a reconciliation between historical documentation, successive modifications and the actual situation.

- ▷ **Decision support in normal or degraded situations**. AI does not make decisions, but proposes hypotheses, options or trade-offs, while explaining their consequences. In **industrial crisis management**, it enables **comparison of scenarios** (stabilise or shut down, continue or isolate).
- ▷ **Assistance with communication in multilingual and multi-profile environments**. The objective is to support cooperation, reduce misunderstandings and rapidly align the different stakeholders. To this end, AI can act as a **technical and cultural translator**, not merely a linguistic one, which can become invaluable on international worksites (instantaneous, context-aware translation of occupational and safety terminology), with heterogeneous teams (operators, engineers, subcontractors), while also rephrasing critical messages according to the recipient's profile. Some applications concern crisis cells, by providing summaries from multiple sources and in multiple languages.

2.5 The 'monitored and secured system' petal

Predictive maintenance: AI learns the "normal behavior" and detects subtle deviations that are not accessible to human perception **[Azeta et al. 2026]**.

- ▷ There are already numerous applications in refineries, power plants, and chemical plants **applied to rotating machinery** (pumps, turbines, compressors): sensors continuously measure vibration, temperature, pressure, and acoustics, with concrete uses such as detecting a bearing that is slowly wearing out, anticipating a shaft failure or seizure, and scheduling a shutdown before a sudden breakdown.
- ▷ Other applications concern intermittent critical equipment (safety valves, relief valves), where **AI correlates sparse data** (opening/closing, micro-leaks) with historical data to identify a valve that is malfunctioning without ever being tested under real operating conditions, with the aim of reducing intrusive tests.

Real-time monitoring to "see what the operator cannot see" in order to reduce individual alarms and increase intelligent "system-level" alerts. AI-based process anomaly detection relies on the simultaneous monitoring of hundreds of parameters (temperatures, flow rates, pressures, compositions) to identify abnormal combinations (rather than an isolated value) and raise an alert before a thermal runaway or chemical drift. Another example: AI uses micro-variations in pressure, acoustic signals, and thermal imaging to detect a hydrogen or ammonia leak before it can be detected by odorization or threshold-based sensors **[Zaidi et al. 2026]**.

AI for the cybersecurity of industrial systems. AI learns the "normal" traffic of industrial networks and detects unusual commands, inconsistent sequences, and any suspicious activity on programmable logic controllers **[Aslam et al. 2025]**. However, the expected benefits could also have a downside, by "opening up" AI-based industrial systems to new cyberattack vulnerabilities. The Gartner consulting firm recalls that by 2028, custom AI applications could account for **50% of cybersecurity incident response efforts** in enterprises, compared with less than 5% in 2024. According to Gartner forecasts, these systems, often deployed without sufficient security testing, will be much more difficult to secure over the long term[3].

Computer vision for hazardous areas. This connects with the "human protected from risks" petal, with the direct objective of reducing serious and fatal accidents through intelligent cameras that automatically detect human presence in restricted areas, failure to wear PPE, or visible leaks (plume, flame, discharge). These systems are already widespread in the oil & gas and heavy chemical industries, as well as at Seveso sites, for monitoring flare stacks, storage tanks, and ATEX zones **[Vukicevic et al. 2024]**.

[3] ICTjournal, 2026, Gartner warns of the security challenges associated with custom AI applications.

2.6 The 'enhanced feedback and learning' petal

The scientific literature converges on one key point: **AI does not improve experience feedback because it "automates" the learning of lessons**, but because it enables rapid, cross-cutting, and systemic analysis of event volumes that are inaccessible to conventional human analysis.

AI to improve the reporting of critical information (beyond formal reporting). In many industries (manufacturing, energy, construction), incident and near-miss reports are very short, subjective, and heterogeneous. They consist of free-text reports. Platforms integrating AI technologies analyze these free-text accounts, comments, and sometimes even informal exchanges (tickets, emails, or internal messages) to detect recurring patterns (fatigue, rule circumvention, or poor coordination), or weak signals that no one had explicitly categorized. AI plays the role here of a tireless reader of thousands of small narratives, capable of noting: *"beware, we are seeing this combination of conditions occur more and more frequently"* [Hashmi et al. 2024].

AI and the production of lessons learned: from storage to actual learning. This corresponds to the transition from incident investigation to the systemic analysis of events:

- ▷ Traditional systems record events; they do not learn.
- ▷ AI-assisted systems, by contrast, search for *patterns*: automatic clustering of similar events, reconstruction of typical degradation chains, and identification of recurring *precursors*.

AI for anticipating critical scenarios (before the accident). Based on lessons learned, AI simulates "plausible futures" in which it can estimate where, when, and under what conditions a degraded scenario is most likely to occur. Some platforms already refer to ***predictive safety*** and ***leading indicators*** derived from historical data to produce a statistical picture of possible futures, built from past errors… including those that "broke nothing" (see the multisector literature review on these possibilities, [Park and Kang 2024]).

AI and positive lessons learned: learning from what worked well. This is still an emerging example, but conceptually a very powerful one. The aim is to identify successful operator adaptations. AI can analyze accounts of events in which a serious problem occurred but was intelligently recovered by the teams. By comparing similar situations, some of which ended badly and others that were successfully recovered, it can highlight effective adaptation strategies, forms of tacit expertise, and operator-system trade-offs that work. Although this aspect remains rare in current industrial practice, it is explicitly discussed in recent approaches to **learning systems** and in critiques of lessons learned processes focused solely on failure (see the FonCSI publications cited above).

3

The macro-level findings of the workshop

This chapter presents the results of the foresight workshop organised by FonCSI in mid-2026. It draws on contributions from participants from a wide range of backgrounds and organisations. Participants were invited to exchange views very freely on the opportunities and difficulties encountered, as well as, from a foresight perspective, on scenarios for the impact of AI deployment on safety.

3.1 Effective deployment of AI for industrial safety is expanding rapidly

By mid-2026, the use of AI for industrial safety applications in the companies represented at the foresight workshop is already evident in at least three areas. In particular, and most widely distributed among the companies represented, are applications of generative AI to safety documentation (design, online use, document simplification). Also, autonomous robotics in high-risk environments. Finally, with growing adoption, intelligent vision-based monitoring systems for hazardous work situations, situations vulnerable to attacks, and maintenance.

The expected safety gains are immense. AI and robotics are extremely effective at eliminating lethal exposure:

- by supporting safer design through their ability to search deeply through available data, propose scenarios of future hazardous situations, and verify design codes;
- by multiplying the value of existing operating experience, in an almost automatic manner;
- by greatly improving the early detection of failures and threats in high-risk facilities, particularly by detecting atypical signals;
- by warning operators, through all available alert mechanisms, of proximity hazards in work situations, thereby reducing serious and fatal accidents. This is even more evident when AI makes it possible to replace operators in dangerous situations with autonomous robotics;
- finally, generative AI has considerable potential as an intelligent collaborative tool for online assistance with diagnosis and access to documentation.

However, **most uses remain at an early stage** in the participants' experience: they are more often imagined than observed in the field. This is often accompanied by an admission of not being aware of, and sometimes discovering *a posteriori*, the actual penetration of AI applications within departments, without any associated risk analysis, given the breadth of their diffusion, which may even be systemic at the scale of the company. The same applies to effects that are often feared, but for which evidence is still lacking, and which are still expressed in conditional and aspirational terms.

A **consumer use that is simultaneously penetrating the enterprise**: several of the companies represented at this foresight workshop do not yet have knowledge of applications for safety in the field that have been "officially" validated. However, all acknowledge the existence of **off-label use**, difficult to regulate, of generative AI by operators and departments, with forms of assistance that are more or less official: in offices, in the field, sometimes using personal computing equipment, and even in design departments, where the use and tracking of AI as a coding aid is still poorly regulated. It is therefore understandable that the link between

these general uses that are already present within the company, with limited oversight, and industrial safety is more indirect but potentially real, and requires a degree of caution.

3.2 Validating well-known threats, alongside some original concerns

3.2.1 Some concerns shared by the industrial community

The discussions during the foresight workshop noted and agreed with many of the concerns regarding the introduction of AI technologies in industry settings that are well-identified, though generally expressed in conditional terms, in the academic and technical literature. This literature has already been summarized in recent FonCSI publications **[Marsden and Steyer 2025; Le Coze and Antonsen 2023; Bieder et al. 2026]**.

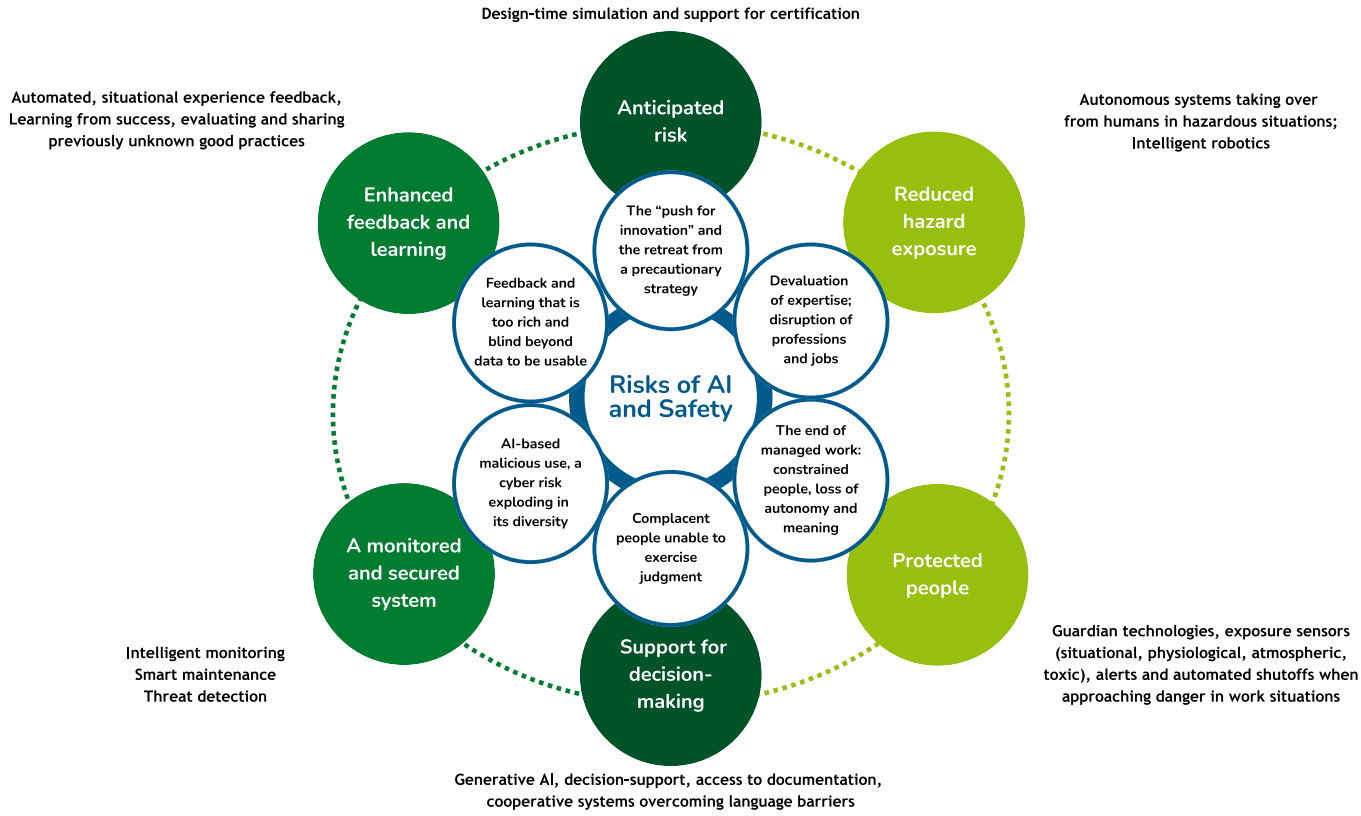

Figure 3.1 *The main concerns regarding industrial safety generated by the introduction of AI technologies in high-hazard industry sectors.*

We retain from these findings four broad generic concerns, represented in figure 3.1:

1. A **difficulty in understanding the machine's reasoning and in cooperating with it**: the new learning capabilities of AI-based systems, with a considerable "black box" effect, have no equivalent in the technological past. This ushers in an era of very high complexity and new risks in mutual understanding and cooperation between humans and autonomous systems. This is the area best described in the literature (see the cited FonCSI publications).
2. A **loss of skills**, with a feared erosion of critical skills (***skill fade***), coupled with an effect of "distance from reality" and overconfidence, and with the gradual disappearance of the tacit knowledge of expert front-line operators.
3. A **difficulty in regaining control**, compounded by the previous point, with a reduced capacity to diagnose a degraded situation when automation fails and a risk of blindly accepting a "statistically high-performing" solution.
4. A **risk of new cyber vulnerabilities** capable of exploiting biases and technical blind spots. AI learns from past situations and may fail to detect new scenarios, whether intentionally or unintentionally introduced into critical vulnerabilities. This risk is related to the emergence

of new failure modes, involving attacks or data manipulation, as well as, more broadly, dependence on vulnerable digital infrastructures.

These four broad generic concerns are supplemented by other **more systemic and organisational concerns**, linked to the shift in the social/societal model induced by the penetration of these technologies:

- Concerns regarding **job losses and occupational transitions**. Building on the model of creative destruction and the rebound in employment proposed by Joseph Schumpeter (early twentieth century **[Schumpeter 1935, 1951]**), one might expect the introduction of AI technologies to eliminate jobs during a first phase, but subsequently to generate new jobs in an equivalent or greater number. However, for the first time, this rebound model is being called into question by the advent of AI. The issue is becoming a major concern for corporate human resources (HR) departments **[Aghion et al. 2025; Kogelmann 2025; Rizvi 2026]**, as many authors believe that AI-related job creation **does not compensate for the destruction of career trajectories**: gains are concentrated in the owners of capital rather than on labour. Contrary to the Schumpeterian model, **the benefits of AI are delayed, concentrated and conditional** on substantial organisational reforms. Paradoxically, these same HR departments are already facing a need for new skills in safety and AI technologies, which are rarely available today.
- An **organisational risk** involving the familiar "illusion of safety" ("the robot handles it"), resulting in a weakening of the culture of operational vigilance and mistrust and, *in fine*, underinvestment in human training.
- A potential **significant negative impact on quality of working life** for workers who retain their jobs, with **surveillance perceived as social control**. Cameras and sensors may be experienced as 1984-like surveillance, and trust may deteriorate if the stated purpose is not clearly the protection of workers. AI risks flattening the meaning of work. A good operational feedback process is also narrative, contextual and contradictory, whereas AI will tend to flatten complexity into "clean" categories, with a risk of normative recentralisation. AI may become a tool that reinforces compliance at the expense of intelligent local adaptations.
- A **shift in (legal) responsibility, linked to the opacity of models**. Decisions become difficult to explain after an incident, raising questions of responsibility and liability (frontline operator, operating company, system integrator, or AI provider).

3.2.2 Some original contributions from the workshop

Participants in the foresight workshop produced a list of additional concerns arising from observations of the development of AI within their companies.

- It will be necessary **clearly to specify when and what we are referring to when discussing safety risks associated with AI**: are we referring to the operator, the customer, the system, the country, and over what time horizon? One of the major transformations brought about by the introduction of AI technologies remains the contraction of space and time and the inherently global approach to its deployment, both within companies and across society.
- Some companies are beginning to purchase **off-the-shelf AI-based systems developed "elsewhere"**. This raises questions about the suitability of a learning system trained in a different environment and country from our own, and therefore from that of the company, both in technical and cultural terms. In the same vein, participants observe that certain AI applications (employee surveillance by camera), already widely implemented in some countries, will have greater difficulty becoming established in France given cultural sensitivty and privacy protection laws.
- **The greatest difficulty in regaining control is probably still to come, with a generational gap**: as long as experts with frontline experience remain present, the system will continue to function, but when these skills disappear, who will still be able to verify what the AI proposes?
- In all the companies represented, **AI deployment is promoted and/or led by a dedicated unit reporting to the highest level of the company**. This explicit organisational structure is important because it may explain a somewhat dual attitude among participants:

they are concerned, or even involved ("it is a company-wide project"), but do not feel ownership ("there are people specifically responsible for it" or "it is bigger than me").

- More generally, **change management for the company and for safety is still in its early stages**. While the local gains from introducing an AI-based system are a key factor driving local adoption within departments, the company's "major systemic adaptation" is not yet genuinely a priority. It will become one, particularly given the risk of "dehumanising" safety.
- Following on from the previous point, participants note that their **human resources departments are not ready**. More generally, AI is more than a simple tool: it will disrupt organisations and work groups. HR departments will then be on the front line on these issues, particularly with regard to industrial relations, even though considerable uncertainty remains and the agenda of their current challenges may limit their capacity to think about the near future shaped by AI.
- An even **greater impact of AI on managerial positions**: beyond job losses and a hypothetical "creative destruction" process, the workshop drew a strong distinction between the impact on white-collar and blue-collar workers. It highlighted that, for the first time in recent history, job losses could be greater among white-collar workers.
- A **social acceptance of AI that raises major questions about "how to manage" adaptation**, with a pace of innovation exceeding all conventional mechanisms of industrial adaptation, both within and outside the company.

3.3 What scenario for the impact of AI deployment on safety?

The workshop continued by attempting to trace the long-term trajectory of the frequency of industrial incidents/accidents, following the gradual diffusion of AI technologies within companies and society. This discussion was very prospective in nature.

3.3.1 Two known trajectories as points of comparison

The workshop was launched with a presentation of two well-known trajectories describing the impact of new technologies on accident rates:

- That of **automation in aviation**, which is relatively straightforward and can be divided into three phases: an initial increase in accident rates, with considerable difficulties in the retraining of pilots; lessons learned from accidents which led to technological improvements, together with generational renewal among pilots but no job losses; and, finally, a slow but continuous improvement in safety that became established over the long term.

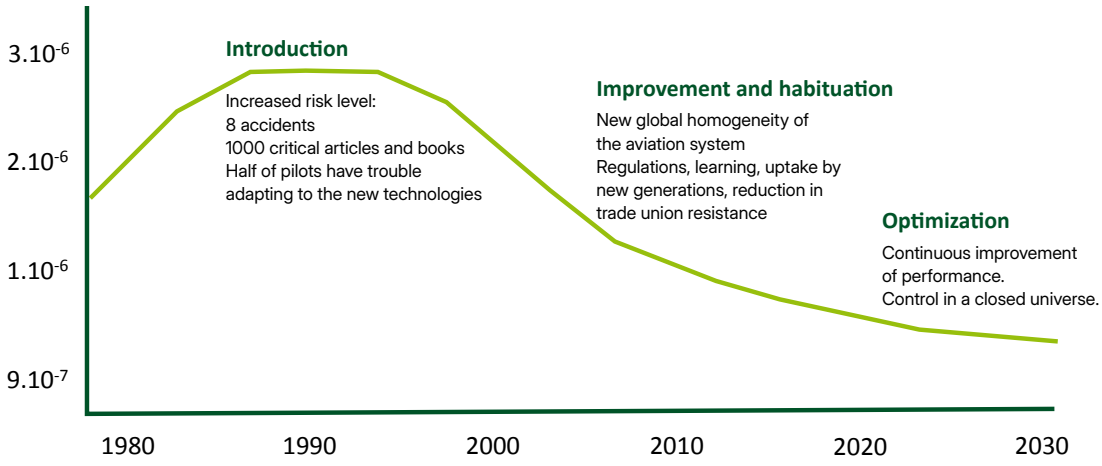

Figure 3.2 Schematic trajectory of accident rates following the introduction of automation in aviation during the 1980s–2000s.

▷ That of the **introduction of digital technologies** (Digital 2.0 and subsequently 3.0 in the enterprise): a more complex and initially difficult evolution, with a highly fragmented range of offerings often driven by consultants and coupled with a generational brake. This was followed by adoption, largely encouraged by parallel adoption among the general public, and then by deployment on a very large scale, generally without job losses and even with net job creation. However, it encountered two initially unexpected obstacles: dependence, resulting from the lack of control over a sovereign cloud infrastructure, and cyber risk, which intensified and diversified over time.

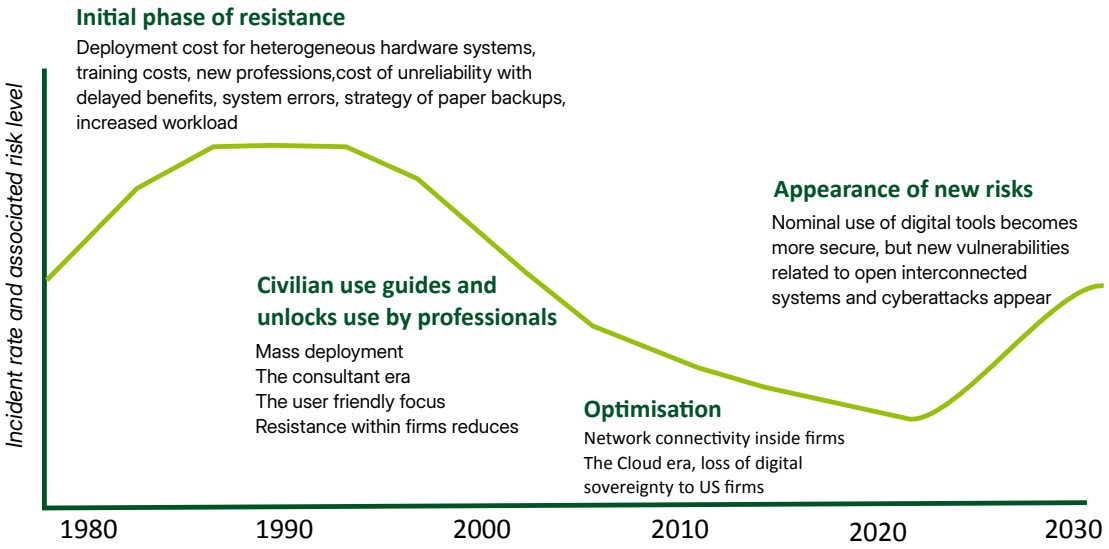

Figure 3.3 *Schematic trajectory of accident rates following the introduction of digital technologies in industrial settings.*

3.3.2 Evolutionary phases for AI technologies as discussed at the foresight workshop

This section reflects the contributions of participants in the foresight workshop concerning possible trajectories for accident rates in high-hazard industrial systems following the introduction of AI technologies. Participants were invited to discuss, using a very open-ended foresight approach, possible scenarios for this evolution, as compared with the trajectories seen for automation in aviation and digital technologies in industry. The participants were organized in five groups, whose contributions are discussed below.

Scenario proposed by the first group:

An initial period of continuous improvement in accident rates should be observed; however, secondary effects could gradually outweigh the expected benefits, triggering a political and economic turning point marked by a massive reduction in the use of AI, amounting to a form of "**AI winter**".

At the same time, a steady increase in the number of AI systems would also be observed, leading to a continuous intensification of performance. This dynamic, initially viewed favourably by the company (and sometimes even actively pursued), **could generate tensions, or even a social backlash** or a loss of social acceptance from the perspective of workers. This could then lead to a reduction in the deployment of these technologies, or at least a limitation of their areas of application, in order to prevent an uncontrollable employment crisis. Finally, the learning capabilities of AI systems, approaching forms of creative behaviour and distinct from those of current systems, would be expected to result in stabilisation, or even a reduction, in the incident rate; however, a risk of overconfidence and loss of control, rare but abrupt, could also emerge.

Scenario proposed by the second group:

Two trends would coexist, one of which has already begun:

- A negative trend associated with the **anticipated increase in psychosocial risks**, resulting from work intensification, the reduction of value-added tasks, and a loss of bearings, with an impact primarily on white-collar workers, even more so than on blue-collar workers. A new social reality, characterised by differing adaptive capacities among individuals, would also have to be addressed.
- A positive trend concerning the risks of accidents affecting people and industrial facilities could be observed, with an initial plateau in accident rates, followed by a **significant reduction in risks**; a third phase of rebound could then occur, insofar as, once the effectiveness of AI had been established, its increased integration into processes would lead to a renewed increase in risks.

Scenario proposed by the third group:

A rapid improvement in the accident-rate curve should initially be observed. However, systems would subsequently be discovered and deployed in a cascade across multiple applications, leading to a more fluctuating accident-rate curve, with highly variable outcomes depending on the type of safety considered. In all cases, **a short time scale would be assumed**, on the order of months or years rather than decades.

In this context, **regulatory intervention** could be observed and could potentially prevent an increase in accident rates following accidents. At the same time, increasing complexity would be observed, requiring greater learning efforts.

Finally, although this is by no means certain, a phenomenon of uncontrolled and malicious runaway behaviour could occur, resembling situations described in science fiction.

Scenario proposed by the fourth group:

Time accelerates with AI. We move from a timescale measured in decades to one measured in years. **AI should affect the local stability of systems differently** depending on the form of deployment chosen: either deployment will allow autonomous systems to develop, in which case the short-term evolution will probably involve an increase in risks; or deployment will be based on operators remaining in control, in which case a reduction in accident rates would appear immediately, even if the *quick wins* were likely to fade over time.

However, the systemic dimension would emerge as the most concerning: unlike sectors such as aviation, the framework would extend beyond the enterprise to affect the organisation of society as a whole, opening the way to contrasting scenarios of evolution.

In this context, an unstable environment associated with the introduction of AI could lead to a **world profoundly reshaped by its effects**, both in terms of forms of enterprise and regulation. New types of accidents could then occur, **including a risk of a global "crash"** (rapidly extending beyond the boundaries of the enterprise and potentially affecting industrial sectors, or even states), with a cascading effect.

Scenario proposed by the fifth group:

This group limited the scope of its reflections to generative AI.

At the present stage, **society is almost ahead of the enterprise**: citizens are ready to accept this transformation, and there is a convergence of interest between individuals and businesses, as was the case with the arrival of digital technologies among the general public.

In this context, a positive effect should initially be observed; however, this effect could gradually diminish as a result of a loss of human expertise. If the accident-rate curve were then to rise as a result of accidents, a counter-reaction by companies could be observed, with the reintroduction of experts. At the same time, career trajectories would be reconsidered and more cautious limitations on use could be introduced through regulation.

Furthermore, another dimension of risk would be identified, this time at the geopolitical level, affecting society as a whole as well as the loss of skilled employment. Regulatory and political developments could indeed occur, but this counter-reaction would take place in a

context of already ongoing reductions in public-sector staffing and deregulation, with public administrations themselves facing increasing complexity. Overall, one would observe the **growth of a societal risk** that is ineffective in addressing problems, while companies would remain relatively more agile in adapting at their own level.

Finally, there is considerable uncertainty regarding the pace of AI progress: current capabilities can be compared with those of five years ago, with already considerable differences; **any projection ten years into the future appears particularly uncertain**.

4

Conclusion

The main conclusions from this foresight workshop on AI applied to industrial safety were broadly aligned with the existing literature. In this respect, **its content is not revolutionary**.

Nevertheless, **seven important features** can be identified, which reflect current concerns and are somewhat more original when compared with the conventional discourse on the subject:

1. **AI is already present in the enterprise**, diffuse and subject to varying degrees of support and oversight, with a rate of penetration measured in months and years rather than decades.
2. The arrival of AI (particularly generative AI) has no precedent in the past in terms of the immediacy and scale of its effects, which are systemic and global from the outset because of its deep penetration into all areas of societal use. The resulting **social upheaval** (employment, uses, organisation and governance of the enterprise) may ultimately be more consequential than the risk of conventional industrial accidents.
3. Some companies have upstream working groups or strategic initiatives, but the **practical implications of the changes that need to be made are not yet clear**, particularly for HR departments, which are on the front line in dealing with the social impact of the arrival of AI technologies.
4. All these concerns about AI point to a broadly shared consensus: **humans and the enterprise must retain control of this industrial revolution**... but all participants have some doubts about whether this will be the case in practice as deployment proceeds.
5. The safety benefit of introducing AI is **potentially immense**, and should be particularly substantial in reducing serious and fatal accidents.
6. However, this benefit will also encounter a **massive secondary effect**: a loss of meaning, disengagement on the part of human workers, and, for the first time, an impact that will probably be greater on white-collar workers than on blue-collar workers.
7. The first accidents involving AI-based systems will provide the impetus for **legislation and new recommendations**. The regulatory trajectory could then significantly reduce the use of these systems (and lead to an "AI winter").

Bibliography

Aghion, P., Bunel, S., Jaravel, X. *et al.* (2025). *How different uses of AI shape labor demand: Evidence from France*. AEA Papers and Proceedings, 115:62–67. DOI: 10.1257/pandp.20251047.

Aslam, M. M., Tufail, A., Gul, H. *et al.* (2025). *Artificial intelligence for secure and sustainable industrial control systems - a survey of challenges and solutions*. Artificial Intelligence Review, 58(11). DOI: 10.1007/s10462-025-11320-9.

Azeta, J., Omeche, T. T., Daniyan, I. *et al.* (2026). *Artificial intelligence and robotics in predictive maintenance: a comprehensive review*. Frontiers in Mechanical Engineering, 11. DOI: 10.3389/fmech.2025.1722114.

Bieder, C., Amalberti, R., Pariès, J. *et al.* (2024). *Industrial safety in the age of "living with": Uncertainty, complexity and rising expectations*. Cahier de la sécurité industrielle 2024-02, Fondation pour une culture de sécurité industrielle. www.foncsi.org, DOI: 10.57071/420lwu.

Bieder, C., Kamaté, C., Laroche, H. *et al.*, Ed. (2026). *Living with the New Safety Landscape I: Rethinking Safety Management in the Context of Cascading Uncertainties*. SpringerBriefs in Safety Management. Springer. ISBN: 978-3032293350.

Gihleb, R., Giuntella, O., Stella, L. *et al.* (2022). *Industrial robots, workers' safety, and health*. Labour Economics, 78. DOI: 10.1016/j.labeco.2022.102205.

Hashmi, F., Hassan, M. U., Zubair, M. U. *et al.* (2024). *Near-miss detection metrics: An approach to enable sensing technologies for proactive construction safety management*. Buildings, 14(4). DOI: 10.3390/buildings14041005.

Huber, J., Anzengruber-Tanase, B., Schobesberger, M. *et al.* (2025). *Evaluating user safety aspects of AI-based systems in industrial occupational safety: A critical review of research literature*. International Journal of Environmental Research and Public Health, 22(5). DOI: 10.3390/ijerph22050705.

ILO (2025). *Revolutionizing health and safety: The role of AI and digitalization at work*. Technical report, International Labour Organization. DOI: 10.54394/KNZE0733.

Kogelmann, B. (2025). *Creative destruction and the autonomous life*. Journal of Business Ethics, 197(4):659–671. DOI: 10.1007/s10551-024-05721-z.

Le Coze, J.-C. and Antonsen, S., Ed. (2023). *Safety in the Digital Age: Sociotechnical Perspectives on Algorithms and Machine Learning*. SpringerBriefs in Safety Management. Springer. ISBN: 978-3031326332.

Lopez, P., Erwin, J., Hopper, D. *et al.* (2025). *From traditional robotic deployments towards assisted robotic deployments in nuclear decommissioning*. Frontiers in Robotics and AI, 12. DOI: 10.3389/frobt.2025.1432845.

Malenfer, M., Héry, M. and Clerté, J. (2022). *Intelligence artificielle au service de la santé et sécurité au travail : Enjeux et perspectives à l'horizon 2035*. Hygiène & sécurité du travail, 269:87–96. www.inrs.fr/media.html?refINRS=VP%2036.

Marsden, E. and Steyer, V. (2025). *L'IA et la gestion de la sécurité : enjeux et questions clés*. Cahier de la Sécurité Industrielle 2025-01, Fondation pour une culture de sécurité industrielle (FonCSI). www.foncsi.org, DOI: 10.57071/iae289.

NSC (2023). *Improving workplace safety with robotics*. Technical report, US National Safety Council.

Park, J. and Kang, D. (2024). *Artificial intelligence and smart technologies in safety management: a comprehensive analysis across multiple industries*. Applied Sciences, 14(24). DOI: 10.3390/app142411934.

Rizvi, S. A. H. (2026). *The productivity paradox revisited: Artificial intelligence, labour market restructuring, and the inequality trap*. Oxford Journal of student scholarship. DOI: 10.65161/rec4D7lEPE4FTkgap.

Sakshi, G., Shreya, J., Vidya, S. *et al.* (2024). *Enhancing worker safety in the construction industry through smart helmet technology, a review*. International Journal of Creative Research, 12(10). www.ijcrt.org/papers/IJCRT2410466.pdf.

Schumpeter, J. A. (1935). *Théorie de l'évolution économique : Recherches sur le profit, le crédit, l'intérêt et le cycle de la conjoncture*. Dalloz. ISBN: 978-2247004547.

Schumpeter, J. A. (1951). *Capitalisme, socialisme et démocratie*. Payot. ISBN: 978-2228883177.

Vukicevic, A. M., Petrovic, M., Milosevic, P. *et al.* (2024). *A systematic review of computer vision-based personal protective equipment compliance in industry practice: advancements, challenges and future directions*. Artificial Intelligence Review, 57(12). DOI: 10.1007/s10462-024-10978-x.

Zaidi, S. H. H., Shenfield, A., Zhang, H. *et al.* (2026). *A systematic review of anomaly and fault detection using machine learning for industrial machinery*. Algorithms, 19(2). DOI: 10.3390/a19020108.

You can extract these bibliographic entries in BibTeX format by clicking on the paperclip icon.

Reproducing this document

FonCSI supports **open access** to research results. For this reason, we distribute the documents that we produce under a licence that allows sharing and adaptation of the content, as long as credit is given to the author following standard citation practices.

With the exception of the FonCSI logo and other logos and images it contains, this document is licensed according to the Creative Commons Attribution licence. You are free to:

- ▷ **Share**: copy and redistribute the content in any medium or format;
- ▷ **Adapt**: remix, transform, and build upon the material for any purpose, even commercially.

as long as you respect the **Attribution** term: you must give appropriate credit to the author(s) by following standard citation practices, clearly indicating the title and URL of the original work, and provide a link to the license.

Translations: If this work is translated, the following disclaimer must be added along with the attribution: This translation was not created by FonCSI, which is not responsible for the content or accuracy of this translation.

Adaptations: In case of an adaptation of this work, the following disclaimer must be added along with the attribution: This is an adaptation of an original work published by FonCSI. Responsibility for the views and opinions expressed in the adaptation rests solely with the author or authors of the adaptation and are not endorsed by FonCSI.

You can download this document, and others in the *Cahiers de la Sécurité Industrielle* collection, from FonCSI's web site.

FONCSI
Fondation pour une culture de sécurité industrielle

Foundation for an Industrial Safety Culture
a public interest research foundation

www.FonCSI.org

6 allée Émile Monso – CS 22760
31077 Toulouse cedex 4
France

Email: contact@FonCSI.org

ISSN 2100-3874

FONCSI

Fondation pour une culture
de sécurité industrielle

6 allée Émile Monso
ZAC du Palays - CS 22760
31077 Toulouse cedex 4

www.foncsi.org