Proceedings of the 35th European Safety and Reliability & the 33rd Society for Risk Analysis Europe Conference
Edited by Eirik Bjorheim Abrahamsen, Terje Aven, Frederic Bouder, Roger Flage, Marja Ylönen

doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P4508-cd

AI-Driven Safety Systems: Reducing Risk in Complex Workplaces and High-Stakes Task

Francesco Di Paco

Department of Civil and Industrial Engineering, University of Pisa, Italy. E-mail: francesco.dipaco@phd.unipi.it

Luca Burattini

Department of Engineering, University of Perugia, Italy. E-mail: luca.burattini@collaboratori.unipg.it

Luciano Di Donato

Department of Technological Innovations and Safety of Plants, Products and Anthropic Settlements, INAIL, Italy. E-mail: l.didonato@inail.it

Roberto Gabbrielli

Department of Civil and Industrial Engineering, University of Pisa, Italy. E-mail: roberto.gabbrielli@unipi.it

Luca Landi

Department of Engineering, University of Perugia, Italy. E-mail: luca.landi@unipg.it

Francesco Marcelloni

Department of Information Engineering, University of Pisa, Italy. E-mail: francesco.marcelloni@unipi.it

Leonardo Marrazzini

Department of Civil and Industrial Engineering, University of Pisa, Italy. E-mail: leonardo.marrazzini@unipi.it

Marco Palumbo

Department of Information Engineering, University of Pisa, Italy. E-mail: marco.palumbo@ing.unipi.it

Marco Pirozzi

Department of Technological Innovations and Safety of Plants, Products and Anthropic Settlements, INAIL, Italy. E-mail: m.pirozzi@inail.it

Complex workspaces that involve workers, machines, and tools often harbor residual risks due to both intended and unintended interactions among these elements. For instance, many incidents arise from operator misuse of machinery, such as working with tampered safeguards or completely removed protections. Additionally, dangerous situations can occur when machines operate without specific auxiliary devices, increasing potential risks within the workspace. The state of the art only provides written warnings in the use and maintenance manual for most of these risks and no more efficient technical solutions have been proposed. In this study, a prototypal machine assembly, comprising a robot and a multimodal lathe, is utilized to assess and mitigate risks in a complex workplace. This workspace is divided into different pre-defined zones depending on the task to be completed and the presence of operators. A comprehensive risk assessment is conducted before and after integrating IoT sensors, such as RFID tags and Computer Vision, integrated with an Artificial Intelligence (AI) module with standard safety systems that comply with the Machine Directive. This approach offers promising solutions to mitigate the consequences of operator errors and potential machine malfunctions. The paper explains how well-known hazards in workspace can be significantly reduced by strategically deploying sensors to monitor specific tasks. Multiple sensors are employed to oversee the case under examination, ensuring a redundant check with sensors based on different technologies to reduce Common Cause Failures of those innovative systems. By combining traditional safety measures with advanced sensor technologies and AI, it is possible to enhance the overall safety of complex workspaces. The proposed system not only addresses common hazardous situations but also provides a proactive approach to managing risks, ultimately contributing to a safer working environment.

Keywords: Artificial Intelligence, Machinery Safety, Errors, Computer Vision, RFID.

1. Introduction

In January 2027, the new EU Machinery Regulation (EU) 2023/1230 (MR) (‘Regulation 2023/1230‘) will come into effect, replacing the current Machinery Directive (MD) (‘Directive 2006/42/CE Machinery Directive). Throughout this paper, these legislative acts will be referred to as MR and MD. This regulation introduces significant changes in machine safety standards. Of relevance to this work are provisions permitting the use of Artificial Intelligence (AI)-based technologies to ensure safety functions and enabling human-machine coexistence in collaborative working environments (commonly referred to as collaborative robots).

Currently, some systems on the market implement safety functions for human protection using devices that are not certified according to a specific Type C product standard. Examples include advancements in anti-collision systems for industrial robots and autonomous vehicles, as well as personnel detection solutions utilizing RFID technology.

In today's industrial landscape, minimizing operator risks associated with moving parts typically relies on protection systems and physical barriers, such as safety barriers or guards. However, certain tasks related to the machinery's lifecycle - such as troubleshooting during malfunctions, maintenance, or mitigating the risk of tampering - are often managed solely through the third step of risk reduction measures, which involve warning signs or operating instructions. With the rise of Industry 4.0 (Condry 2023), the integration of advanced sensors and systems is becoming increasingly common. From a standardization perspective, these are addressed by the technical specification ‘IEC TS 62998-1:2019’. This specification outlines the requirements for designing and integrating safety-related sensors and safety-related sensor systems aimed at protecting individuals. It focuses particularly on systematic capabilities and applies to scenarios where safety technologies are not covered by specific regulatory standards.

The technologies previously mentioned, such as collaborative robots, have seen increased adoption in recent years (Keshvarparast et al. 2024), particularly for tasks aimed at minimizing psychological and physical stress on operators during repetitive work. Under intended use conditions, efforts must focus on eliminating or minimizing discomfort and fatigue as much as possible, adhering to ergonomic principles. These principles, already established in the MD, have been expanded to address challenges posed by the human-machine interface, especially with machines exhibiting self-evolving behavior. In the coming years, the demand for new, safe machinery systems — comprising machine tools (e.g., machining centers, lathes) and collaborative robots operating in shared workspaces with minimal or no physical guarding — will grow. This trend reflects the increasing necessity for true collaborative work environments. The new requirements and methods designed to protect operators will include:

- Detecting potentially hazardous objects, such as tools inadvertently left on moving parts after maintenance or improperly placed machine components (e.g., swarf or chip containers).
- Identifying the removal or tampering of essential guards.
- Monitoring entry into or exit from hazardous zones during operation.
- Employing machine learning and artificial intelligence to manage safety functions and address emerging risks.

In the following sections, we will present the practical potential for risk mitigation in complex production environments through a prototype of a small production system currently in the testing phase. The discussion will begin with an overview of the prototype and a risk analysis demonstrating how these innovative, redundant systems can effectively reduce risks. This will be followed by an explanation of the AI algorithms and system communication protocols employed to enable real-time risk reduction.

2. Complex workplace and risk analysis

The production system under analysis comprises a multimodal CNC lathe and a UR3e robot, which work collaboratively to handle and machine a workpiece. Figure 1 illustrates the layout of the machines and the positioning of walls within the hypothetical workroom. This initial setup is crucial for defining the workspace, which is divided into three distinct zones based on the operator's position relative to the equipment. Zone 1 encompasses the area adjacent to the lathe. Zone 2 includes the region surrounding the workpiece holders, both before and after the turning process, as well as the circular area swept by the robot during its movements. Zone 3 represents the space where the operator can interact with and control the machine assembly via the command console. This zoning helps to clarify the interactions between the operator, the lathe, and the robot within the workspace.

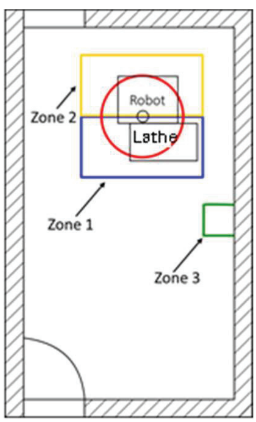

Fig. 1. Workspace surrounding the prototype.

The risk assessment for the described workspace is conducted using the well-established "bottom-up" method ('ISO/TR 14121-2:2012'). The process begins by defining all possible operational states of the system, followed by identifying the associated risks for each state. In the case of the robot's activity, two operational states are considered: full speed and reduced speed. However, since the consequences of impacts at reduced speed have not yet been analyzed, the robot is categorized in this phase as either stationary or moving. During the design phase, criteria must be established to determine when movement is enabled or disabled. To address this, an enable signal has been incorporated into the system:

For the CNC lathe, four operational modes, as described in the machine's user manual, are defined:

- Manual Mode: This mode permits manual processing. The door guard may be opened, while other guards remain in place. It requires the involvement of a simple lathe operator, an experienced operator, or a tool technician.
- Automatic Mode: This mode enables automatic processing through CNC programs. The spindle guard may be opened, while other guards remain in place. Unlike manual mode, this mode allows for inexperienced operators, provided they are limited to part replacement tasks while the machine is stopped.
- Setup Mode: This mode is used for preparing the machine for processing. The carriage guard must be in place, but other guards may be opened or removed to facilitate setup. It requires the involvement of a simple lathe operator, experienced operator, tool technician, or maintenance operator.
- Service Mode: This mode is designed for cleaning and lubricating the machine under safe conditions. Guards are not required, and only a maintenance operator is permitted to use this mode.

These operational modes are crucial for identifying hazards and performing a thorough risk assessment. While many risks are mitigated using traditional safety functions in compliance with standards such as 'ISO 13849-1:2015' and the MD, some hazards remain unaddressed by the current state of the art. These hazards, relevant to the case under study, are presented as "Use Cases" (UC) in Table 1. Typically, these situations are categorized as residual risks, with mitigation limited to providing information to the operator. However, the new EU MR introduces opportunities to leverage innovative technologies for advanced risk mitigation. The following section outlines the proposed supervisory system, and its key components designed to address these challenges and achieve enhanced safety outcomes.

2.1. IoT Sensors Description

This section outlines the IoT sensors incorporated into the proposed prototype and their roles in the experimental supervisory system. The preliminary design includes the following elements:

- Wearable RFID Devices: Operators wear RFID tags integrated into their work coveralls or personal protective equipment. These tags are identified and tracked by a network of antennas placed around the machine, allowing the system to locate operators and associated objects within the workspace.
- Computer Vision (CV) Module: the spatial configuration of objects in the workplace is measured by detecting ArUco markers attached to the parts of production systems that need to be monitored. The reliability of this approach is supported by a detection accuracy close to 1 in well-illuminated environments, as reported in the literature (Hu, DeTone, and Malisiewicz 2019). AI models for pattern recognition, such as those implemented in the OpenCV library (Bradski 2000) for detecting ArUco markers and the RT-DETR model (Zhao et al. 2024) for detecting operators, are used to process images captured by cameras properly positioned in the workplace.
- Safety-Related Software: Data from the wearable devices and cameras are processed by safety software, enabling a supervisory operator to monitor the system's functionality. In compliance with point 1.2.1, Annex III of Regulation (EU), the software can intervene, particularly in emergencies, to ensure safety through corrective actions.

To explore the benefits of these advanced technologies in the workplace, a preliminary system design has been developed. Based on earlier considerations, this framework emphasizes the critical role of integrating innovative technologies. Autonomous devices equipped with AI modules and computational capabilities play a key role in processing data and calculating the Performance Level (PL) of the safety functions.

While determining the PL of the supervisory system poses challenges, segregating sensor outputs provides a robust foundation. Each sensor's data is analyzed independently before being transmitted to the AI module responsible for decision-making. This integration of IoT sensors and AI systems allows machines to interpret new operational scenarios, detect potential risks to operators, and issue commands, such as stopping machine operations when necessary. The integrated system, illustrated in Figure 2, shows interactions between its components. Using a combination of RFID, CV, and AI technologies, it can identify hazardous conditions, such as faults, malfunctions, or improper actions, based on the machine's operational status. When risks are detected, the supervisory system communicates with the machine control system to ensure operator safety, for example, by triggering an emergency shutdown. This approach underscores the transformative potential of IoT and AI in creating smarter and safer work environments.

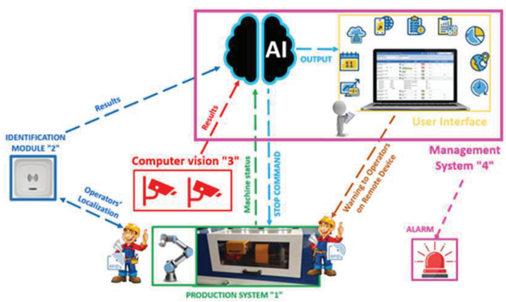

Fig. 2. Prototype scheme components and their relationship.

2.1. Description of AI Module

The AI module was developed as a Rule-Based Expert System (Tan 2017), designed to store knowledge obtained from domain experts and the Machinery Regulation and to make decisions when specific conditions are met. The architecture of the Expert System, illustrated in Figure 3, comprises three main components:

- Knowledge Base (KB): The KB stores the system's knowledge using an ontology to model the workspace. This ontology defines relevant entities, their properties, relationships, and a set of rules that guide decision-making.

- Inference Engine (IE): The IE processes input data by applying the appropriate rules from the KB. It executes these rules to produce output responses tailored to the detected conditions.
- System Interface (SI): The SI serves as the user interface, enabling non-expert users or other application programs to interact with the system. Through the SI, users can input new data and access the system's output.

The structure of the AI module is highly versatile. The KB can be developed in a scalable manner, adapting to the specific physical system under consideration. The integration of a real machine tool and a cobot ensures a thorough analysis of human-machine interaction while keeping the number of rules at a manageable level. In more complex scenarios, such as real industrial environments, the KB can be expanded accordingly with additional rules, while the hardware components will scale in parallel to accommodate the increased system complexity. The development of the AI module begins with defining a formal representation of the workspace using an ontology. This ontology captures the relevant entities in the system as classes, along with their instances, properties, and relationships. Table 1 provides an example of the lathe's representation within the defined ontology, highlighting how the workspace is formally modeled to enable accurate decision-making. This structured approach ensures that the AI module effectively integrates domain knowledge, regulatory requirements, and real-time data to support advanced safety functions.

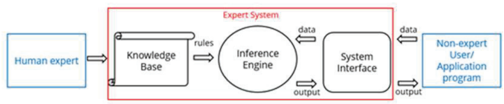

Fig. 3. Expert System architecture

Table 1. Representation of the lathe according to the defined ontology.

ENTITY	CLASSES	PROPERTIES	PROPERTY VALUES	INSTANCES
Lathe	Lathe	has_guards	List of Guard instances	lathe
		power	"ON", "OFF"	
		operative_mode	"MANUAL", "AUTO", "SETUP", "SERVICE"	
	Guard	presence	"DETECTED", "UNDETECTED"	-
		position	"MOUNTED", "UNMOUNTED"	
		is_removed	True, False	
	UnmovableGuard	is_subclass_of	Guard	cart_guard tailstock_guard
	MovableGuard	is_subclass_of	Guard	door_guard spindle_guard
		state	"OPEN", "CLOSED"	
		is_tampered	True. False	

After creating the ontology, the full set of rules is defined, with each rule following the structure: IF condition THEN action. These rules are designed based on risk analysis, which identifies and describes all reasonably foreseeable hazardous situations. Each rule condition specifies a unique combination of property values that could result in such situations. For example, when the lathe is powered on (power = "ON") and operating in automatic mode (operative_mode = "AUTO"), the door guard must remain closed. To enforce this, the system includes a rule to verify this condition. If the condition is not met, a "door guard tampering" warning is triggered:

IF lathe.power == "ON" AND
lathe.operative_mode == "AUTO" AND
door_guard.state == "OPEN"
THEN door_guard.is_tampered = True

During runtime, the AI module receives new data from the IoT sensors. The IE evaluates the rules, activating those whose conditions match the incoming data. The module responds by either sending a stop signal to the production systems or by notifying an alarm to the operators.

The redundancy of the system significantly enhances the overall reliability of the apparatus. Specifically, the integration of multiple detection mechanisms results in a more robust hypothesis testing framework. By combining CV and RFID, it is possible to simultaneously improve both specificity and sensitivity, thereby reducing erroneous state diagnoses For instance, in the context of door tampering detection, the safe state of the door guard (state A) is verified through both CV and RFID, leading to a diagnosis (B: door guard closed) with a joint probability P(A∩B), which is higher than the probability achieved by either system alone (enhanced specificity). Similarly, an unsafe state (A-, door guard not safe) is detected by both CV and RFID, yielding a corresponding diagnosis (B-, door guard not closed) with a joint probability P(A−∩B−), thus improving overall sensitivity. Since an incorrect diagnosis would require both systems to fail simultaneously, the combined system exhibits greater sensitivity compared to individual detection methods.

3. Complex workplace and risk analysis

Defining specific risk scenarios is the crucial first step. For each identified UC, one or more sensors are assigned to monitor specific parameters, with their monitoring frequency determined by the sensors' technological capabilities. Equally critical is establishing the type of response to be triggered when a hazardous situation is detected. Table 1 summarizes these concepts, highlighting the role of various sensors in addressing each scenario.

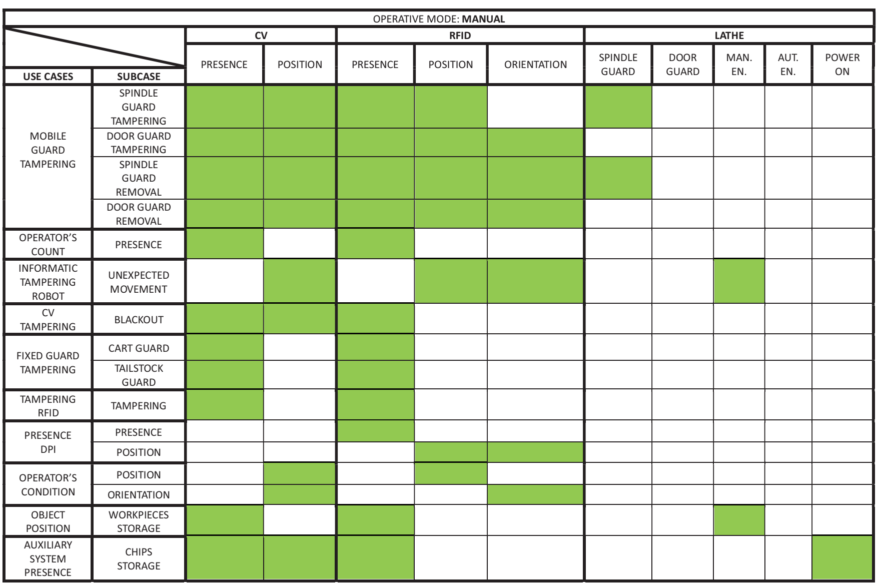

Fig. 4. Use Cases for the manual operational state of the lathe

Figure 4 illustrates the role of each IoT sensor for the specified UC. Since the Numerical Control of the lathe provides four operational modes (Manual, Automatic, Setup, and Service), each with distinct safety conditions, additional matrices like the one presented have been developed to accommodate these variations. Defining these UCs and the corresponding technologies enabled the identification of a set of I/O signals required for communication between the AI system and the other prototype's components, including CV, RFID, and the production system. The primary goal is to develop a prototype that functions as a conventional machine assembly during standard tasks while providing enhanced safety operations when activated by the responsible operator. This ensures the system complies with the MD guidelines and prevents certain operations, such as the robot's collaborative mode, from being engaged when the supervisory system is inactive. To activate the supervisory system, the operator must turn the mode selector on the control console to "Enable AI". This signal is sent to the AI module, which then checks the IoT sensor statuses to determine if supervision can begin. If the conditions are met, the AI module sends an "AI-Ready" signal back to the lathe terminal using Normally Closed logic. If this signal is interrupted or if the two output channels of the AI module do not match, the production system will stop following the supervisory system's safety instructions. Any signal fault is immediately detected by the lathe's safety PLC, ensuring a prompt response to prevent potential hazards.

As shown in Figure 5(b), the electric circuits of the supervisor and production systems are isolated using relays to avoid electrical issues, such as short circuits, from affecting both systems. The safety signal is then transmitted from the lathe to the robot's safety terminal, highlighted in yellow, ensuring seamless integration of the safety protocols across all components.

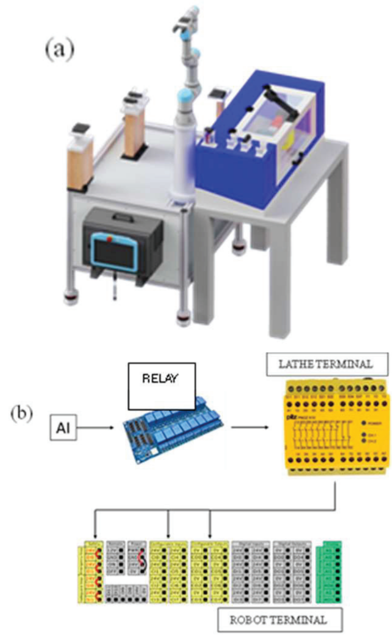

Fig. 5. (a) Rendering of the machine assembly, (b) safety signal path.

For details on stop function categories, refer to 'EN 60204-1:2018'.

When a hazard concerns only the robot, the lathe remains unaffected, allowing its work cycle to continue uninterrupted. In such cases, a Category 2 stop is particularly valuable, as it halts the robot without disrupting the current cycle phase of the lathe. While a Category 0 stop offers a higher level of safety by removing power, it may prevent the resumption of the work process from the last phase, potentially resulting in the loss of a workpiece. This is especially critical in processes that can take several hours or involve high-value workpieces. The safety signal ultimately reaches the robot terminal to initiate the corresponding safety response. As outlined in the user manual, Universal Robots e-Series robots are equipped with integrated safety functions and I/Os designed to connect with other machines and additional protective devices. These safety functions and I/Os are implemented in compliance with ISO 13849, achieving Performance Level d with Category 3 architecture. The robot also supports three types of stops, as defined by the EN 60204-1 standard (Emergency, safety, limit violation and anomaly detection) ('EN 60204-1:2018'). The authors suggest that this prototype system could serve as a foundation for further discussion on integrating innovative sensors and AI algorithms into machinery, aligning with the new MR set to take effect in 2027.

4. Conclusions

The next steps in this study involve constructing and testing the prototype to evaluate its ability to ensure operator safety, particularly in addressing currently underexplored risks, such as the presence of auxiliary systems and tampering. Automating the workpiece turning cycle and enabling the transfer of safety signals will require a well-regulated and efficient architecture between the machines. Additionally, auxiliary components must be designed to securely hold the workpiece before and after the robot's pick-and-place operations. To validate the system's effectiveness, rigorous testing will be conducted under typical risk scenarios involving the prototype's turning activities. These tests aim to provide a comprehensive evaluation of the safety mechanisms, ensuring the system reliably protects operators in real-world conditions. The integration of IoT sensors, AI-driven supervision, and advanced safety protocols is expected to significantly enhance workplace safety, addressing both traditional hazards and those that remain challenging for current solutions. Ultimately, this approach lays the foundation for smarter and safer industrial environments, fostering a more symbiotic relationship between humans and machines.

Acknowledgement

This research is an integral part of the ongoing: "AISAFETY – Sistema smart integrato basato sull'intelligenza artificiale per la gestione della sicurezza degli operatori in processi di produzione – BRiC 2022 ID 40 Project, co-founded by INAIL" by Bando Ricerche in Collaborazione (BRiC – 2022) – Activity Research Plan 2022-2024.

References

Bradski, G. 2000. 'The OpenCV Library'. *Dr. Dobb's Journal of Software Tools*.

Condry, M. W. 2023. 'Technologies of Industry 4.0 and Opportunities They Create for Product Innovation'. *IEEE Engineering Management Review* 51 (4): 185–90.

'Directive - 2006/42 - EN - Machinery Directive - EUR-Lex'. https://eur-lex.europa.eu/eli/dir/2006/42/oj/eng.

'EN 60204-1:2018 - Safety of Machinery - Electrical Equipment of Machines - Part 1: General Requirements'.

Hu, D., D. DeTone, and T. Malisiewicz. 2019. 'Deep ChArUco: Dark ChArUco Marker Pose Estimation'. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8428–36. Long Beach, CA, USA: IEEE.

'IEC TS 62998-1:2019. Safety of Machinery - Safety-Related Sensors Used for the Protection of Persons.'

'ISO 13849-1:2015. Safety of Machinery - Safety-Related Parts of Control Systems - Part 1: General Principles for Design.' ISO.

'ISO/TR 14121-2:2012. Safety of Machinery — Risk Assessment Part 2: Practical Guidance and Examples of Methods.' ISO.

Keshvarparast, A., D. Battini, O. Battaia, and A. Pirayesh. 2024. 'Collaborative Robots in Manufacturing and Assembly Systems: Literature Review and Future Research Agenda'. *Journal of Intelligent Manufacturing* 35 (5): 2065–2118.

'Regulation - 2023/1230 - EN - EUR-Lex'. https://eur-lex.europa.eu/eli/reg/2023/1230/oj/eng.

Tan, H. 2017. 'A Brief History and Technical Review of the Expert System Research'. *IOP Conference Series: Materials Science and Engineering* 242 (1): 012111.

Zhao, Y., W. Lv, S. Xu, J. Wei, G. Wang, Q. Dang, Y. Liu, and J. Chen. 2024. 'DETRs Beat YOLOs on Real-Time Object Detection'. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 16965–74. Seattle, WA, USA: IEEE.