Taxonomy for Early Safety Analysis of Human-Robot Interaction

Meenakshi Manjunath[1], Philipp Kranz[1], Bastian Tenbergen[2] and Marian Daun[1]

[1]Center for Robotics, Technical University of Applied Sciences Würzburg–Schweinfurt, Schweinfurt, Germany
[2]Department of Computer Science, State University of New York at Oswego, Oswego, United States

Abstract

In the era of automation and digitization, not every task can be automated in an economically feasible manner. Considering industrial settings, where robots often work alongside and with shop-floor workers, safety risks and threats often confine these robots to simpler and caged tasks. This setup introduces safety challenges, as robots must not only perform tasks efficiently, but also ensure the physical safety of nearby humans. In practice, safety aspects are often considered late in development or addressed through isolated technical measures. This paper presents a safety taxonomy that supports the identification of relevant hazards early in the development process. The taxonomy links safety hazards in production environments to their potential causes, detection methods, and mitigation strategies, taking into account specific process characteristics. It supports safety engineers and developers in analyzing risks during the design phase and implementing appropriate detection and mitigation measures.

Keywords

Human-Robot Interaction, Safety Analysis, Safety Model

1. Introduction

The growing digitization and interconnectivity of manufacturing systems—driven by advancements in industrial automation—has led to the increasing adoption of collaborative robots on the shop floor. These robots work in close physical proximity to human operators and assist in shared tasks such as handling workpieces, executing production steps, or providing support during assembly processes [1, 2]. This setup, often referred to as human-robot interaction, introduces a new class of safety challenges.

Unlike traditional industrial robotics, which rely on physical barriers or two-hand actuation mechanisms to prevent human injury [3], collaborative environments remove such constraints in favor of increased production flexibility and efficiency [4]. As a result, robots (i.e, collaborative robots) and humans share workspaces, overlap in task execution, and even engage in direct physical interaction. While this transition offers clear productivity benefits, it also exposes human operators to new risks, particularly due to the force, speed, and torque capabilities of modern robotic systems [5].

Designing robot behavior to default to conservative safety responses—such as stopping upon unexpected proximity or contact—can significantly hinder productivity and limit the effectiveness of human-robot interaction [6]. At the same time, reactive approaches that address safety issues only after incidents occur are insufficient, as they fail to prevent serious harm. Thus, both the robot control software and the collaborative production process must be certifiably safe before deployment begins [7].

This paper presents a safety taxonomy aimed at supporting the early stages of developing interactive robotic systems. The taxonomy classifies typical safety hazards in production environments along with their possible causes, detection methods, and mitigation strategies. With the overview it provides, safety issues can be anticipated early in development and addressed before they lead to hazardous situations during operation. This enables humans and robots to work together efficiently while maintaining a high level of safety.

ER2025: Companion Proceedings of the 44th International Conference on Conceptual Modeling: Industrial Track, ER Forum, 8th SCME, Doctoral Consortium, Tutorials, Project Exhibitions, Posters and Demos, October 20-23, 2025, Poitiers, France
0009-0005-6421-1450 (M. Manjunath); 0000-0002-1057-4273 (P. Kranz); 0000-0002-0145-4800 (B. Tenbergen); 0000-0002-9156-9731 (M. Daun)

This paper is structured as follows. Section 2 presents the theoretical underpinnings and the rationale behind the proposed safety taxonomy. Section 3 proposes the safety taxonomy, and Section 4 provides an initial evaluation of our proposed taxonomy. Finally, Section 5 concludes this paper.

2. Background and Related Work

2.1. Types of Human-Robot Interaction

Human-robot interaction (HRI) encompasses different interaction modes between human operators and robotic systems. In industrial contexts, these interactions must be carefully planned and managed to ensure both operational efficiency and operator safety. To better understand the risks and safety requirements associated with different degrees of proximity and cooperation, HRI in industrial production is often categorized into five distinct types [8, 9]. These categories reflect increasing levels of physical and task-based interaction, each with its own safety implications:

- **Cell Operation:** The robot operates within a fenced-off cell, physically separated from human operators. There is no direct interaction, and tasks are executed independently. This form of interaction resembles the base level of automation—analogous to ”no autonomy” in autonomous driving taxonomies [10]—and offers the highest level of physical separation and inherent safety.
- **Coexistence:** Humans and robots operate in the same general environment without physical barriers, but in distinct workspaces and without coordinated interaction. Safety is maintained through spatial separation and awareness, with each party performing tasks independently.
- **Synchronized Interaction:** Human and robot share the same workspace, but perform their tasks at different times. For example, a robot may prepare parts that are subsequently assembled by a human. Although physical proximity exists, direct contact is avoided through sequential task planning.
- **Cooperative Interaction:** Humans and robots simultaneously perform different sub-tasks of a shared process within the same workspace. While they do not physically interact with the same component, their activities must be closely coordinated to prevent interference and ensure safety.
- **Collaborative Interaction:** Human and robot work together on the same component at the same time and within the same workspace. This form of interaction involves the highest degree of physical and temporal overlap, requiring advanced sensing, control strategies, and safety protocols to manage shared actions and avoid accidents.

2.2. Need for Safety Assessment of Human-Robot Interactive Systems

While collaborative robots i.e. cobots are equipped with built-in safety features, these alone are insufficient to guarantee safety across varying production environments. The certification of hardware components does not account for the contextual factors introduced by the specific setup and nature of human-robot interaction. According to standards such as ISO 15066 [11] and VDI-EE 4030 [12], the safety of the entire interactive robotic system must be assessed and assured before deployment in industrial settings.

Particular attention must be paid to the planning complexities that arise from the interaction between human operators and robots in shared workspaces [3], including:

- **Path Planning Complexity:** Robots often operate on pre-programmed trajectories or adapt dynamically using artificial intelligence to respond to changing environments and human behavior [13, 14]. Ensuring safety under these adaptive behaviors adds to the complexity of system planning.
- **Interaction Modalities:** HRI can range from physical collaboration to non-contact coordination. Each mode introduces different types and levels of risk depending on task timing, spatial arrangement, and responsiveness [15, 16].

- **Task Complexity:** Tasks assigned to collaborative robots may involve multiple interdependent steps, require fine precision, or necessitate synchronization with human actions, all of which increase the challenge of ensuring safe execution [15].

2.3. Safety Analysis of Human-Robot Interaction

The development of safety-critical systems across domains such as automotive and aerospace is governed by well-established safety standards. For instance, ISO 26262 [17] in the automotive sector and ARP4761A [18] in aviation mandate structured development processes that begin with qualitative hazard identification—commonly using techniques such as Functional Hazard Analysis (FHA) or Failure Modes and Effects Analysis (FMEA) [19]—followed by quantitative verification of mitigation strategies using methods such as Fault Tree Analysis or Markov models [20].

For industrial automation, a comparable standard is IEC 61508 [21], which governs the functional safety of electrical and programmable systems. However, unlike its counterparts in other domains, IEC 61508 does not prescribe specific techniques for hazard identification. As HRI introduces a complex interaction paradigm into industrial production, this gap becomes particularly relevant. In traditional hazard analysis, engineers rely heavily on domain expertise and precedent to identify potential safety issues [20]. However, this approach is limited in HRI scenarios, where real-world examples are sc, and and the human operator introduces unpredictability. While the robot's behavior can be formally defined and analyzed, human actions are inherently variable and often deviate from intended procedures in ways that are difficult to model in advance.

To address some of these challenges, process-oriented approaches such as Leveson's Systems-Theoretic Process Analysis (STPA) have been proposed [22, 23]. Unlike traditional techniques, STPA emphasizes the control structure and causal scenarios within a system, making it suitable for analyzing complex, dynamic processes such as those found in collaborative production settings. Once hazards have been identified, mitigation typically involves two dimensions. First, direct hazards arising from the robot's mechanical behavior—such as uncontrolled arm motion or unsecured object handling—can often be addressed through specific safety tasks or design changes [24, 25]. Second, indirect hazards that stem from interaction complexity—such as unpredictable human behavior, task dependencies, or environmental variability—require runtime strategies involving real-time monitoring and perception, often using sensing technologies like 3D cameras [7].

In our previous work, we proposed the use of model-based engineering approaches to support early safety analysis in HRI systems [26, 27]. These approaches utilize goal modeling techniques to systematically identify potential safety hazards at a conceptual stage—particularly those arising from robot task execution and the complex dependencies involved in HRC [28]. In addition, we explored the use of these models to develop digital twins of robotic systems that can serve as a runtime safety system [29].

2.4. Existing Taxonomies

Taxonomies have proven to be effective tools for structuring safety assessments across various safety-critical domains. In aviation, for instance, hazards are systematically categorized with a strong emphasis on mechanical reliability and human factors [30]. Similarly, in the context of embedded and cyber-physical systems, existing taxonomies address the interplay between hardware and software components and their associated failure modes [31]. Many of these approaches also incorporate software-related safety risks, as discussed in foundational work on system safety engineering [20]. However, these existing taxonomies were developed for systems where human involvement is either indirect or tightly constrained, and they do not directly translate to the unique characteristics of HRI. In HRI systems, humans and robots operate in shared physical workspaces and engage in dynamic or task-dependent interactions. This tight coupling between human behavior and robot control logic introduces specific risks that are not sufficiently captured by general-purpose taxonomies.

Although robots used in such interactive environments undergo rigorous hardware-level certification

[11, 12], such certification does not account for safety in the overall system context. The integration of a certified cobot into a dynamic environment with human operators introduces new sources of hazards that need to be addressed at the system level.

Several efforts have been made to categorize safety considerations specific to HRI. For example, Akalin et al. [32] identify three major dimensions of safety in collaborative settings: *physical safety*, which includes the use of physical or behavioral safeguards to prevent harm; *perceived safety*, which relates to how humans interpret the robot's behavior and judge the environment as safe or unsafe; and *data security*, which involves ensuring the confidentiality and integrity of information exchanged between the robot and its environment.

Another work emphasizes task structure as a basis for understanding and managing HRC safety. Taxonomies that classify interactions based on human and robot task roles [33] provide useful abstractions for decomposing complex collaborative tasks into smaller, manageable components. Such decomposition can help reveal potential points of failure or unsafe dependencies in the workflow.

Complementing this, the OCRA framework [34], which extends the robot knowledge representation platform KnowRob [35], offers a logic-based approach to modeling and managing task execution in HRC settings.

3. Safety Taxonomy for Human-Robot Interaction

3.1. Overview of Safety Hazards

To systematically identify, analyze, and mitigate safety hazards in HRI, it is essential to classify hazards based on their origin within the system. Previous work, such as by Berx et al. [36], emphasizes the importance of origin-based classification. Building on this idea, we propose a taxonomy that distinguishes hazards arising from the human operator, the robot, and the interaction between them, in addition with detection and mitigation strategies for these hazards.

Robot-specific tasks often require basic but critical safety measures. For instance, verifying the correct closure of a gripper before a lifting action is essential to prevent object drops and resulting injuries [24, 25]. In such cases, continuous monitoring through optical, tactile, or proximity sensors plays a key role in enabling safe behavior. Sensor-based monitoring serves not only as a reactive safety mechanism but also as a proactive tool for hazard prevention, particularly in dynamic environments.

Beyond the robotic subsystem, safety hazards frequently arise from human-robot dependencies. Previous analyses have identified specific types of hazards that put human operators at risk in interactive and collaborative workflows [26]. These include, for example, coordination mismatches, timing delays, or unclear task boundaries between humans and robots.

To better structure these risks, hazards in HRI can be grouped into three general categories [37, 7]:

- **Physical hazards**, such as collisions or unintended contact, which require persistent spatial monitoring and compliance with dynamic safety zones.
- **Ergonomic hazards**, including physical strain caused by repetitive tasks or awkward postures, which call for assessment of work performance and task design to reduce operator fatigue.
- **Coordination and communication hazards**, which arise when task execution between humans and robots is not synchronized, potentially leading to misunderstandings or unsafe transitions.

Beyond single human–robot interactions, more complex configurations—such as one human interacting with multiple robots or collaborating human–robot teams—pose additional safety challenges, particularly concerning coordination between systems [38]. Identifying hazards based on their source—whether they stem from hardware, software, interaction patterns, or system-level coordination—allows for more targeted mitigation strategies and clearer traceability throughout the design and validation process.

3.2. Safety Hazard Classification

A HRI system can be broadly divided into two principal components: the robot and the human operator. In most cases, the robot involved in such scenarios is a collaborative robot, commonly referred to as a cobot. Accordingly, safety hazards are grouped into two primary categories: human-related hazards and system-specific hazards.

1. **Human-Related Hazards** cover physical injuries as well as physiological and ergonomic strain experienced by human operators. These hazards emerge from the physical proximity, working in the same workspace and task-sharing characteristic of HRI systems. Table 1 provides an overview of the human related safety hazards.

    - ***Collision, Crushing and Trapping***: These include direct physical hazards such as collisions, crushing, or impact-related injuries. Since HRI typically takes place in a shared, unguarded workspace, even minor deviations in human or cobot behavior can result in severe incidents. For example, collisions may occur due to unexpected human movement, dropped components during pick-and-place operations, or uncontrolled cobot motion stemming from mechanical or software faults.
    - ***Psychological Strain***: Strain refers to the cognitive and physiological fatigue experienced by human operators due to prolonged periods of concentration or sustained workload. In HRI, while cobots perform tasks tirelessly, human operator may experience stress and reduced performance over time. This form of strain can negatively impact decision-making, task execution speed, and overall mental well-being.
    - ***Physical Ergonomics***: This category focuses on the physical impact of repetitive tasks, awkward postures, or physically demanding movements. Unlike robots, human bodies are prone to discomfort and musculoskeletal disorders resulting from poorly designed workspaces or excessive repetition. Such ergonomic challenges may lead to long-term injuries and decreased productivity by the human operator.

2. **System-Specific Hazards** refer to failures and limitations within the cobot or its environment that can compromise safety during collaboration. Table 2 provides an overview of the different system specific hazards.

    - ***Cobot Malfunctions***: These include electrical or mechanical failures such as short-circuiting, power surges, or actuator faults. Such malfunctions can disrupt intended operations and lead to unsafe cobot behavior, including uncontrolled movements or dropped payloads.
    - ***Synchronization Errors***: These arise from misaligned timing or coordination between human and cobot actions. In shared workflows, such failures can result in process inefficiencies, assembly errors, or even physical accidents if one party moves before the other has completed a dependent task.
    - ***Sensor Errors***: These affect the cobot's ability to perceive its environment accurately. Failures in tactile, optical, or proximity sensors may cause incorrect object detection, pose estimation, or distance calculation. This can result in mishandling parts, incorrect force application, or unintended proximity to human operators.

3.3. Safety Hazards, Detection Methods, and Mitigation Strategies

Only preliminary identification of hazards is not enough to implement safety strategies; it is important to incorporate the right detection method and mitigation strategy as well. Once hazards have been identified, addressing them—particularly those stemming from the mechanical functionalities of cobots—requires the definition of concrete safety tasks. For example, verifying the proper closure of a gripper is critical to prevent dropped objects and resulting injuries.

Safety Hazards Subclass	Hazards	Detection Method	Mitigation Strategies
Collision	• Human error: Unexpected movements or mistakes	• Optical distance sensors • Proximity sensor • Force torque sensor •Tactile sensors	• Training and workshops • Implementation post-process button • Emergency stop response mechanisms
Crushing	• Cobot error: Malfunctions or incorrect response		
Trapping	• Part impact: Dropped or mishandled components		
Psychological Impact	• Mental Fatigue from prolonged periods of concentration • Stress due to physical pain	• Wearable sensors • Tracking work efficiency and product throughput • Tracking of vitals of the operator	• Work efficiency monitoring • Ergonomic redesign of the workspace • Adjust task allocation • Periodic rest between tasks
Physical Ergonomics	• Pain and discomfort due to repetitive motions • Musculoskeletal disorders due to awkward postures		

Table 1
Overview of Human Related Safety Hazards in Human-Robot Interaction

Table 1 and 2 provide a comprehensive overviews of the hazards, the detection method and the mitigation strategies. *Safety hazard subclass* lists the main hazard types. Followed by *hazards*, which provides instances in which the hazard occurs. *Detection methods* specify the relevant detection methods used to identify the hazard. And finally, *mitigation strategies* list corresponding mitigation strategies that are aligned with industrial safety standards.

Collision, Crushing and Trapping (Table 1), i.e. physical hazards, typically arise from sensor misjudgments, unanticipated human movements, or mishandling of components during shared tasks in the workspace.

- **Detection**: These hazards are typically detected using optical distance sensors, force-torque sensors, and tactile sensors, which enable real-time monitoring of proximity, force thresholds, and contact events.
- **Mitigation**: Mitigation strategies include safety training for human operators, the integration of post-process buttons to signal task completion before cobot action, and the implementation of emergency stop mechanisms to immediately halt the cobot in hazardous situations.

Comparative to collision, crushing and trapping; **psychological strain** and **ergonomic strain** (Table 1) pose greater long term hazards for the human operator. Psychological strain may result from prolonged concentration or task repetition, often leading to mental fatigue and reduced performance. Physical ergonomic strain, on the other hand, can result in discomfort, pain, or longer-term musculoskeletal disorders due to repetitive motions or non-optimal postures.

- **Detection**: Strain can be identified through wearable sensors, continuous monitoring of operator vitals, and analysis of task efficiency and throughput metrics.

- **Mitigation**: Appropriate mitigation measures include ergonomic redesign of the workspace, modified task allocation between the cobot and human, scheduling of regular rest breaks, and adjusting task performance parameters to manage workload and reduce operator stress.

Safety Hazards Subclass	Hazards	Detection Method	Mitigation Strategies
Cobot Malfunctions	Short circuiting or power surges	Voltage sensors	Routine maintenance of the cobot
Synchronization Errors	Failed coordination between cobot and human	External camera monitoring system	Reprogramming and recalibration of cobot
Sensor Errors	• Gripper not closing correctly • Incorrect placement of parts • Incorrect proximity calculation	• Tactile sensors • Machine vision system • Proximity sensor • Optical distance sensor	Routine maintenance of the sensors on the cobot

Table 2
Overview of System Specific Hazards in Human-Robot Interaction

System specific threats can be further classified into cobot malfunctions, synchronization errors, and sensor-related failures, as outlined in Section 3.2. These hazards originate from the hardware, software, or sensing components of the cobot, which is illustrated in Table 2, that provides an overview of these hazards along with respective detection methods and mitigation strategies.

Cobot malfunctions refer to failures originating from the electrical or mechanical components of the robot. These can include power surges, short circuits, or actuator-level faults. If not addressed through regular maintenance, such issues may escalate into more severe hazards, including uncontrolled behavior or, in extreme cases, electrical or fire-related incidents.

- **Detection**: These failures are typically detected using voltage sensors that monitor electrical anomalies in the system.
- **Mitigation**: Preventative strategies include routine maintenance to ensure stable cobot operation and prevent unanticipated failures.

Synchronization errors arise when the timing between human and cobot actions is not properly aligned. These errors often occur when the cobot completes its task either faster or slower than the human operator, leading to disruptions in the shared workflow.

- **Detection**: These issues are typically identified through external monitoring systems, often using computer vision to track and evaluate the alignment of human and cobot actions.
- **Mitigation**: Effective measures include reprogramming and recalibration of the cobot, as well as ensuring regular software updates to maintain compatibility with updated workflow.

Sensor errors occur when perception systems on the cobot fail to accurately interpret their surroundings. Such failures can result in incorrect gripper closure, misplacement of parts, or miscalculation of safe distances between the cobot and the human operator.

- **Detection**: These hazards are typically detected using a combination of tactile sensors, proximity sensors, optical distance sensors, and machine vision systems.
- **Mitigation**: The standard mitigation approach involves routine inspection and calibration of sensor systems to ensure reliable and accurate perception during collaborative operation.

Safety Hazard Category	Safety Hazard Subclass	Classical Cell Operation	Coexistence	Synchronized	Cooperation	Collaboration
Human Related Hazards	*Collision*	⇩	▬	⬆	⬆	⬆
	Crushing	⇩	▬	⬆	⬆	⬆
	Trapping	⇩	▬	⬆	⬆	⬆
	Psychological Impact	⇩	▬	▬	⬆	⬆
	Physical Ergonomics	⇩	▬	▬	⬆	⬆
System Specific Hazards	*Cobot Malfunctions*	⬆	⬆	⬆	⬆	⬆
	Synchronization Errors	⇩	⇩	▬	▬	⬆
	Sensor Errors	⇩	⇩	▬	⬆	⬆

Legend	
⬆	High Liklihood of Safety Hazard
▬	Moderate Likelihood of Safety Hazard
⇩	Low Likliehood of Safety Hazard

Table 3
Safety Hazard Probability Across Different Types of Human-Robot Interaction

3.4. Relation between Modalities of Human-Robot Interaction and Safety

Table 3 offers a comparative summary of the likelihood of safety hazards across the different types of HRI. The hazards are organized under two main categories: *human-related hazards* and *system-specific hazards*, as shown in the *Safety Hazard Category* column. Each subsequent column corresponds to a specific HRI modality. Following ISO 12100, the likelihood or probability of the hazard is defined as the qualitative chance that a hazard occurs in a given HRI modality; whereas severity of the hazard is defined as the gravity of harm to the operator should the hazard occur [39, 40]. Consistent with ISO 12100's qualitative treatment of the probability of occurrence of harm, and common practice in EN ISO 13849-1 [41] risk graphs, we express likelihood on a three-point qualitative scale (*low, moderate, high*) for clarity in Table 3.

- **Cell Operation**: In this mode of HRI, the cobot operates within a physically separated cell, completely isolated from human operators. This strict spatial separation effectively minimizes the likelihood of human-related hazards such as collision, crushing, trapping, psychological stress, and ergonomic strain. However, particularly cobot malfunctions, remain relevant. Technical failures—such as power surges or mechanical faults—are highly likely to occur, even in isolated setups.
- **Coexistence**: Here, the cobot and the human operator share the same broader environment but carry out independent tasks in separate work zones. While there is no direct interaction, the absence of physical barriers increases the chances of accidental contact, raising the likelihood of physical injuries to a moderate level. Similarly, psychological stress and ergonomic strain may occur due to ambient factors such as noise, spatial crowding, or continuous machine presence. While synchronization and sensor-related hazards are moderately likely, the probable occurrence of cobot malfunctions remains high.
- **Synchronized**: In synchronized interaction, the human operator and cobot share the same workspace and alternate tasks in a sequential manner. Despite the lack of simultaneous activity, the shared environment presents a high likelihood of physical hazards such as collisions, trapping, or crushing—particularly when task transitions are poorly timed. The need for precise handover and coordination between agents significantly increases the probability of synchronization errors. Psychological strain, ergonomic strain, synchronization and sensor errors remain at moderate occurrence, due to reliance on accurate timing and perception.
- **Cooperation**: In cooperative interaction, both entities operate concurrently in the same space on related tasks, though not on the same workpiece. The simultaneous working pace of the human

operator and cobot increases the probability of physical injuries substantially. Psychological and ergonomic stressors are also high due nature of continuous working. Additionally, the likelihood of especially sensor errors and cobot malfunctions is high. In contrast, synchronization errors are moderately likely, as the tasks performed by the human operator and cobot are not directly interdependent in terms of timing.
- **Collaboration**: This interaction type involves the highest level of integration, where the operator and cobot work together on the same task and often the same object in real time and shared workspace. Due to the continuous, high-dependency interaction, all hazard types—collision, crushing, trapping, psychological stress, ergonomic strain, cobot malfunctions, synchronization failures, and sensor errors—are at their highest likelihood.

4. Evaluation

To evaluate the applicability of the proposed taxonomy and the corresponding probability levels of safety hazards, a use-case-based approach is adopted. The selected use-case is an industrial scenario involving the collaborative picking and assembling of parts along a production assembly line ([3, 1]).

Its structure makes it well-suited for analysis, as it is representative of many real-world manufacturing processes, which captures the range of spatial and functional relationships that characterize HRI. This can be then adapted to each of the five HRI modalities: cell operation, coexistence, synchronized, cooperation, and collaboration. For each modality, the use-case is adjusted to reflect its specific interaction characteristics (Section 2.1). Each scenario is then used to identify possible human-related and system-specific hazards and the probability of the hazards based on the interaction.

The use-case includes one human operator and a single cobot working on the same workpiece to assemble a part in the production assembly line. The steps of the use-case are as follows:

- The cobot picks up a component from a bin.
- It places the component in the correct position and holds it steady.
- The human operator performs a task on the same part (e.g., attaching, aligning, or inserting another piece).
- After the human operator completes their task, the cobot moves the assembled part to a designated area.

Using the definition from Section 2.1 and the safety hazard tables (Tables 1 and 2), the use-case is adapted to each specific HRI type. This allows for a logical assessment of which human-related and system-specific hazards are likely to occur, along with the probability they may pose for the human operator.

4.1. Classical Cell Operation

As outlined in Section 2.1, classical cell operation refers to an HRI modality in which the robot is physically enclosed, operating in complete isolation from the human operator. There is no shared workspace or direct interaction. Instead, coordination is achieved through indirect mechanisms such as a conveyor belt used for transferring workpieces.

In the adapted use-case, the human operator performs the initial step of the assembly process and places the partially completed workpiece onto a conveyor. The cobot, operating within a safeguarded cell, then picks up the workpiece and carries out a simple task such as repositioning or transferring it. This setup ensures spatial and functional separation between the two entities.

Given the physical isolation, human-related hazards such as collisions, crushing, or trapping (Table 1) are highly unlikely. Similarly, the probability of psychological stress or ergonomic strain is minimal, as the human is neither required to coordinate with the robot nor adapt to its movements.

System-specific hazards (Table 2) are less prominent in this setup. Synchronization errors do not apply, as the tasks are performed sequentially and independently. Sensor errors, such as incorrect proximity or part detection, are also less critical in this context, as the robot operates in a predictable and controlled environment. However, cobot malfunctions—particularly electrical failures or unexpected shutdowns—remain relevant. Even though the operator is not directly at risk, such failures can disrupt the workflow, damage workpieces, or compromise the internal safety of the robot. For this reason, appropriate detection and mitigation strategies, including routine diagnostics and voltage monitoring, are still essential.

4.2. Coexistence

Referring to Section 2.1, coexistence refers to a modality of HRI in which the human operator and the cobot occupy the same general workspace but perform their tasks independently, with no requirement for physical interaction or timing-based coordination. Unlike classical cell operation, the cobot is not enclosed in a protective cage. Instead, a spatial separation is maintained through task design and workstation layout to reduce the likelihood of direct physical contact.

In the adapted use-case, the cobot carries out its pick-and-place operations on one side of the shared area, while the human operator performs partial assembly on the other. Although the tasks are functionally decoupled, the absence of physical barriers introduces a moderate likelihood of contact-based hazards. Collisions, crushing, or trapping incidents may occur due to unexpected operator movements, misinterpretation of spatial boundaries, or cobot path deviations. Such physical injuries can arise when proximity constraints are violated, especially in dynamic production settings.

The continuous presence of an active robot in the same workspace may also contribute to moderate levels of psychological impact. Ergonomic issues can result from working in constrained spaces, while psychological stress may stem from heightened alertness or distraction caused by the cobot's movement. This aligns with previously identified cognitive and physical load factors associated with non-contact but close-proximity interaction.

Cobot malfunctions remain a relevant concern in this modality. Electrical faults, actuator errors, or software-related failures may still occur, and their consequences can extend beyond process disruption if the malfunction causes the cobot to encroach into the human's area unexpectedly. Hence, voltage monitoring and predictive maintenance are essential mitigation measures.

In contrast, synchronization and sensor-related hazards are of low concern in this setup. Since the human and robot execute independent tasks with no temporal interdependence or data exchange, coordination failures and sensing errors are less likely to result in safety-critical outcomes.

4.3. Synchronized

In synchronized interaction, as defined in Section 2.1, the human operator and the cobot share the same workspace and carry out their tasks in a sequential yet coordinated manner. Although their actions are not simultaneous, each agent depends on the timely completion of the other's task to continue the process.

When adapting the use-case to this modality, the cobot initiates the workflow by placing a workpiece onto the shared assembly area. The human operator then performs a partial assembly step before the cobot introduces a second component. Once the operator completes the assembly, the cobot retrieves and transfers the finalized part to a designated location. The sequence demands ongoing monitoring and timely responses, both from the cobot and the human operator.

Given the close physical proximity, the likelihood of physical injuries; such as collisions, crushing, and trapping—is considerably high. These hazards arise when task transitions are mistimed, spatial paths are misaligned, or the cobot resumes movement prematurely. Such incidents are particularly likely in environments where shared access to the same workspace is required during task execution.

The interdependence of the task sequence introduces a moderate risk of synchronization errors. These usually stem from miscommunication through outdated control parameters, or delayed human response. Similarly, sensor-related hazards are moderately likely. Since the cobot must rely on its perception systems to accurately interpret the status of the workpiece and the human's progress. Inaccuracies in object detection, motion prediction, or spatial localization can result in operational misjudgments.

Ergonomics and psychological impact also pose a moderate risk. The human operator must maintain situational awareness and work at a pace aligned with the cobot's actions, which can lead to both cognitive fatigue and physical stress, especially over extended periods.

Cobot malfunctions remain a high concern. Despite operating in a structured environment, failures in software execution, actuator reliability, or power delivery can lead to unintended behavior that disrupts the synchronized workflow and poses safety risks to the human operator. Regular diagnostics and system updates are therefore always essential.

4.4. Cooperation

According to the definition in Section 2.1, cooperative interaction refers to a modality of HRI in which the human operator and the cobot work concurrently within the same workspace toward a common task, yet handle different components or process stages. In the adapted use-case, the cobot continuously delivers multiple workpieces into the shared area, while the human operator performs assembly steps on different components. Both agents work in parallel, within close proximity, but do not act on the same workpiece simultaneously.

This spatial overlap introduces a high likelihood of physical hazards. Since the cobot and the operator are active at the same time in the same environment, the risk of collision, crushing, or trapping is significant—particularly when task boundaries are unclear or movements are misaligned.

Ergonomic and psychological strain are highly relevant. The operator must maintain spatial awareness throughout the process, which can result in physical fatigue due to repeated adjustments in posture or workflow interruptions. The constant presence of the cobot may also contribute to cognitive load or reduced concentration, particularly during longer shifts or under time pressure.

Cobot malfunctions continue to pose a high risk. Failures in actuation, power supply, or control systems may cause unintended movements within the shared space, thereby compromising operator safety. Sensor-related hazards are also likely, as the cobot must detect environmental changes, including human presence and dynamic object positioning. Synchronization errors may occur with moderate likelihood, since some level of coordination is required to ensure uninterrupted and safe parallel task execution.

4.5. Collaboration

The collaborative interaction corresponds to the primary use-case described above, in which the human operator and the cobot work simultaneously on the same workpiece within a shared workspace. As outlined in Section 2.1, this modality demands continuous coordination, joint attention to the same object, and sustained spatial awareness from both agents throughout the task.

Due to the absence of physical separation and the simultaneous manipulation of the same object, this configuration presents the highest likelihood of safety hazards across all categories. Physical hazards—such as collision, crushing, or trapping—are highly probable, as even minor deviations in movement or timing can result in direct contact between the human and the cobot.

Human-related hazards such as psychological stress and ergonomic strain are also prominent. The operator must remain highly attentive to the cobot's movements while executing their own tasks, leading to increased cognitive load and potential stress. Extended periods of such interaction can further contribute to physical fatigue or musculoskeletal discomfort due to posture shifts and constrained working conditions.

Among system-specific hazards, cobot malfunctions present a serious risk. Unexpected behavior—such as abrupt halts, trajectory errors, or force misapplication—can endanger the operator, particularly in the absence of physical barriers. Sensor-related hazards are equally critical in this modality, as the cobot must continuously perceive and interpret human actions, gestures, and workspace dynamics with high accuracy. Synchronization errors are also highly likely, as the task sequence depends on tightly coupled and time-sensitive actions from both agents.

As such, collaborative interaction consistently reflects the highest probability ratings across all safety hazard types in the proposed taxonomy when compared to the other HRI modalities.

Using the proposed taxonomy to examine the adapted versions of the use-case across different HRI modalities revealed clear differences in the types and likelihood of safety hazards. In classical cell operation, where the tasks are completely separated in both space and function, human-related hazards were minimal, while system-specific hazards such as cobot malfunctions remained relevant. In coexistence, although the tasks were still independent, working in the same environment introduced moderate physical and psychological strain. For synchronized and cooperative interactions, the degree of dependency in task execution increased, which raised the likelihood of synchronization and sensor errors. In the collaborative setup, where both agents work closely on the same task, all hazard types—physical, cognitive, and system-specific—were highly relevant. This illustrates how the probability and relevance of hazards change depending on the HRI modality, and how the proposed taxonomy aided in identifying hazards.

5. Conclusion

The transition toward human-centric manufacturing represents a fundamental shift in how industrial environments are designed and operated. In contrast to traditional paradigms where human workers and machines were deliberately separated through physical barriers and strict zoning, modern production systems increasingly involve close, often simultaneous interaction between human operators and robotic systems. Using HRI, a new level of adaptability, flexibility and efficiency has been introduced through Industry 4.0. However, this transformation renders many of the conventional safety strategies inadequate. Physical isolation is no longer a desirable solution in environments where humans and robots are expected to work collaboratively or side by side.

In this paper, a safety taxonomy has been proposed which is specifically tailored to HRI systems. The taxonomy categorizes safety hazards based on whether it is human-related or system-specific hazards; and aligns each with relevant detection methods and mitigation strategies. By applying this taxonomy across different HRI modalities, the taxonomy has demonstrated its practical use. The probability of each hazard was introduced and analyzed, highlighting how risk levels evolve depending on the degree of interaction and shared responsibility between the human and the robot.

To evaluate the taxonomy, a common industrial use-case involving the pick-and-assembly of parts by a human operator and a cobot was used. This base scenario was adapted to each of the five HRI modalities defined in the paper: classical cell operation, coexistence, synchronized, cooperative, and collaborative interaction. For each adapted version, the process steps were examined to determine which hazards—both human-related (e.g., physical injuries, psychological stress, ergonomic strain) and system-specific (e.g., cobot malfunction, synchronization errors, sensor failures)—could plausibly occur.

The proposed taxonomy is a foundational checklist for the different modalities of HRI. It helps identify safety gaps and required safety measures early in the planning phase. For each modality and its associated hazards, it guides the selection of appropriate detection methods and mitigation strategies and provides a clear baseline classification for reference during design and planning. It also enables consistent comparison of hazards across HRI types and supports decisions on engineering controls, sensing, and monitoring.

Through this structured evaluation, the taxonomy proved to be both adaptable and comprehensive. It also revealed clear patterns in hazard relevance and probability, showing that as interaction between human and robot increases, both human-related and system-specific hazards increase in likelihood and impact. The evaluation also demonstrated how to use the taxonomy as a checklist in safety planning: for each modality, it helps trace process steps to hazard categories, select suitable detection methods and mitigations, and record decisions for design reviews. In this way, the taxonomy enables consistent analysis across HRI configurations and supports early safety reasoning during system design.

Declaration on Generative AI

During the preparation of this work, the authors employed the following AI tools: *ChatGPT*, *Grammarly*, and *Oxford English Dictionary AI Assistant* to support paraphrasing and grammar checks. All content was then reviewed and edited by the authors, who take full responsibility for the final version of this work.

References

[1] N. Karnik, U. Bora, K. Bhadri, P. Kadambi, P. Dhatrak, A comprehensive study on current and future trends towards the characteristics and enablers of industry 4.0, Journal of Industrial Information Integration 27 (2022) 100294.
[2] M. Peshkin, J. E. Colgate, Cobots, Industrial Robot: An International Journal 26 (1999) 335–341.
[3] V. Villani, F. Pini, F. Leali, C. Secchi, Survey on human–robot collaboration in industrial settings: Safety, intuitive interfaces and applications, Mechatronics 55 (2018) 248–266.
[4] J. Krüger, T. K. Lien, A. Verl, Cooperation of human and machines in assembly lines, CIRP annals 58 (2009) 628–646.
[5] B. C. Jiang, O. S. Cheng, Design for robotic cell safety, Journal of Manufacturing Systems 9 (1990) 169–175.
[6] C. S. Franklin, E. G. Dominguez, J. D. Fryman, M. L. Lewandowski, Collaborative robotics: New era of human–robot cooperation in the workplace, Journal of Safety Research 74 (2020) 153–160.
[7] S. Robla-Gómez, V. M. Becerra, J. R. Llata, E. Gonzalez-Sarabia, C. Torre-Ferrero, J. Perez-Oria, Working together: A review on safe human-robot collaboration in industrial environments, Ieee Access 5 (2017) 26754–26773.
[8] L. McGirr, Y. Jin, M. Price, A. West, K. van Lopik, V. McKenna, Human robot collaboration: Taxonomy of interaction levels in manufacturing, in: ISR Europe 2022; 54th International Symposium on Robotics, VDE, 2022, pp. 1–8.
[9] W. Bauer, M. Bender, M. Braun, P. Rally, O. Scholtz, Lightweight robots in manual assembly–best to start simply, Frauenhofer-Institut für Arbeitswirtschaft und Organisation IAO, Stuttgart 1 (2016).
[10] O.-R. A. D. O. Committee, Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles, SAE International, 2021.
[11] ISO, Iso/ts 15066:2016 robots and robotic, devices - collaborative robots. publisher = international organization for standardization, 2016.
[12] VDI, Consideration of human reliability in the design of autonomous systems, 2022.
[13] A. Mukovskiy, C. Vassallo, M. Naveau, O. Stasse, P. Souères, M. A. Giese, Adaptive synthesis of dynamically feasible full-body movements for the humanoid robot hrp-2 by flexible combination of learned dynamic movement primitives, Robotics and Autonomous Systems 91 (2017) 270–283.
[14] A. Obaigbena, O. A. Lottu, E. D. Ugwuanyi, B. S. Jacks, E. O. Sodiya, O. D. Daraojimba, Ai and human-robot interaction: A review of recent advances and challenges, GSC Advanced Research and Reviews 18 (2024) 321–330.
[15] R. Gervasi, L. Mastrogiacomo, F. Franceschini, A conceptual framework to evaluate human-robot collaboration, The International Journal of Advanced Manufacturing Technology 108 (2020) 841–865.
[16] S. Haddadin, E. Croft, Physical human–robot interaction, Springer handbook of robotics (2016) 1835–1874.
[17] I. ISO, 26262: Road vehicles-functional safety, International Standard ISO/FDIS 26262 (2011) 53–54.
[18] S.-. Aircraft, S. Dev, S. A. Committee, Guidelines and methods for conducting the safety assessment process on civil airborne systems and equipment, SAE International, 1996.
[19] C. A. Ericson, Hazard analysis techniques for system safety, john wiley&sons, Inc, New Jersey (2005) 24–29.
[20] N. G. Leveson, Safeware: system safety and computers, ACM, 1995.

[21] International Electrotechnical Commission, Functional safety of electrical/electronic/programmable electronic safety-related systems - part 1: General requirements, 2010. URL: https://webstore.iec.ch/publication/5517.

[22] N. G. Leveson, A systems-theoretic approach to safety in software-intensive systems, IEEE Transactions on Dependable and Secure computing 1 (2004) 66–86.

[23] N. G. Leveson, Engineering a safer world: Systems thinking applied to safety, The MIT Press, 2016.

[24] Z. M. Bi, C. Luo, Z. Miao, B. Zhang, W. Zhang, L. Wang, Safety assurance mechanisms of collaborative robotic systems in manufacturing, Robotics and Computer-Integrated Manufacturing 67 (2021) 102022.

[25] A. Vysocky, P. Novak, Human-robot collaboration in industry, MM Science Journal 9 (2016) 903–906.

[26] M. Manjunath, J. Jesus Raja, M. Daun, Early model-based safety analysis for collaborative robotic systems, IEEE Transactions on Automation Science and Engineering (2024).

[27] M. Daun, M. Manjunath, J. Jesus Raja, Safety analysis of human robot collaborations with grl goal models, in: International Conference on Conceptual Modeling, Springer, 2023, pp. 317–333.

[28] J. Jesus Raja, P. Kranz, M. Daun, Comparison of dependencies for human-robot interaction types, in: ER2024: Companion Proceedings of the 43rd International Conference on Conceptual Modeling: ER Forum, Special Topics, Posters and Demos, October 28-31, 2024, Pittsburgh, Pennsylvania, USA, 2024.

[29] J. Jesus Raja, M. Manjunath, M. Daun, Towards a goal-oriented approach for engineering digital twins of robotic systems., in: ENASE, 2024, pp. 466–473.

[30] D. O'Hare, The 'wheel of misfortune': a taxonomic approach to human factors in accident investigation and analysis in aviation and other complex systems, Ergonomics 43 (2000) 2001–2019.

[31] B. Tenbergen, A. C. Sturm, T. Weyer, A hazard taxonomy for embedded and cyber-physical systems, in: Proceedings of the 1st International Workshop on Emerging Ideas and Trends in Engineering of CyberPhysical Systems (EITEC), 2014.

[32] N. Akalin, A. Kiselev, A. Kristoffersson, A. Loutfi, A taxonomy of factors influencing perceived safety in human–robot interaction, International Journal of Social Robotics 15 (2023) 1993–2004.

[33] L. Onnasch, E. Roesler, A taxonomy to structure and analyze human–robot interaction, International Journal of Social Robotics 13 (2021) 833–849.

[34] A. Olivares-Alarcos, S. Foix, S. Borgo, G. Alenyà, Ocra–an ontology for collaborative robotics and adaptation, Computers in Industry 138 (2022) 103627.

[35] M. Tenorth, M. Beetz, Knowrob: A knowledge processing infrastructure for cognition-enabled robots, The International Journal of Robotics Research 32 (2013) 566–590.

[36] N. Berx, W. Decré, I. Morag, P. Chemweno, L. Pintelon, Identification and classification of risk factors for human-robot collaboration from a system-wide perspective, Computers & Industrial Engineering 163 (2022) 107827.

[37] V. De Simone, V. Di Pasquale, V. Giubileo, S. Miranda, Human-robot collaboration: An analysis of worker's performance, Procedia Computer Science 200 (2022) 1540–1549.

[38] M. Askarpour, D. Mandrioli, M. Rossi, F. Vicentini, Formal model of human erroneous behavior for safety analysis in collaborative robotics, Robotics and computer-integrated Manufacturing 57 (2019) 465–476.

[39] I. O. for Standardization, Safety of machinery — general principles for design — risk assessment and risk reduction, 2010.

[40] S. Haddadin, A. Albu-Schäffer, G. Hirzinger, Safe physical human-robot interaction: measurements, analysis and new insights, in: Robotics Research: The 13th International Symposium ISRR, Springer, 2010, pp. 395–407.

[41] C. I. O. for Standardization, Safety of machinery — safety-related parts of control systems — part 1: General principles for design, amendment 1, 2021.