

# Automatic Risk Management for Industrial Control Systems

DISSERTATION

zur Erlangung des akademischen Grades

**Doktor der Technischen Wissenschaften**

eingereicht von

**Dipl.-Ing. Pushparaj Bhosale, MSc.**

Matrikelnummer 12037294

an der Fakultät für Informatik  
der Technischen Universität Wien

Betreuung: Dipl.-Ing. Dr.techn. Wolfgang Kastner

Zweitbetreuung: Dipl.-Ing. Dr.techn. Thilo Sauter

Diese Dissertation haben begutachtet:

---

Prof. Dr. Edgar Weippl

---

Prof. Dr. Norbert Pohlmann

Wien, 28. August 2025

---

Pushparaj Bhosale





# Automatic Risk Management for Industrial Control Systems

DISSERTATION

submitted in partial fulfillment of the requirements for the degree of

**Doktor der Technischen Wissenschaften**

by

**Dipl.-Ing. Pushparaj Bhosale, MSc.**

Registration Number 12037294

to the Faculty of Informatics

at the TU Wien

Advisor: Dipl.-Ing. Dr.techn. Wolfgang Kastner

Second advisor: Dipl.-Ing. Dr.techn. Thilo Sauter

The dissertation has been reviewed by:

---

Prof. Dr. Edgar Weippl

---

Prof. Dr. Norbert Pohlmann

Vienna, August 28, 2025

---

Pushparaj Bhosale



# Erklärung zur Verfassung der Arbeit

Dipl.-Ing. Pushparaj Bhosale, MSc.

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Ich erkläre weiters, dass ich mich generativer KI-Tools lediglich als Hilfsmittel bedient habe und in der vorliegenden Arbeit mein gestalterischer Einfluss überwiegt. Im Anhang „Übersicht verwendeter Hilfsmittel“ habe ich alle generativen KI-Tools gelistet, die verwendet wurden, und angegeben, wo und wie sie verwendet wurden. Für Textpassagen, die ohne substantielle Änderungen übernommen wurden, haben ich jeweils die von mir formulierten Eingaben (Prompts) und die verwendete IT- Anwendung mit ihrem Produktnamen und Versionsnummer/Datum angegeben.

Wien, 28. August 2025

---

Pushparaj Bhosale



# Acknowledgements

I would like to express my deepest gratitude not only to the professionals who have guided and supported me but also to everyone whose contributions have been vital to the success of this dissertation. The collaborative environment and collective effort from all involved have been indispensable to this achievement.

First and foremost, I would like to thank my project guides, Dr. Wolfgang Kastner and Dr. Thilo Sauter, for their invaluable guidance, continuous support, and encouragement throughout the course of this research. Their expertise, insights, and constructive feedback have been instrumental in shaping this work. Without their mentorship, this project would not have been possible.

I am immensely grateful to Siegfried Hollerer, Ali Hosseini, and Mukund Bhole, the members of the umbrella group, for our continuous discussion on practical challenges and improvement in the approach and publication. I would also like to thank the students, faculty members, and industrial partners of SafeSecLab for providing a collaborative and intellectually stimulating environment. The regular faculty meetings helped us to understand the industrial needs and keep the approach practical.

The Scientific Advisory Board meetings, held annually, played a pivotal role in keeping the project on track with academic needs. The thoughtful suggestions and guidance from all the board members have been greatly appreciated. Special thanks to Ulrike Weisz and Ruth Fochtner, whose administrative and organizational support made it all come together. Also, my gratitude goes towards Automation Systems Group for their timely responses to my needs.

Finally, I would like to extend my heartfelt thanks to my parents. My Father, Rajaram Bhosale for his motivation and my mother, Shobhana Bhosale, for her unwavering emotional and mental support. Unfortunately, I recently lost my mother and her absence is deeply felt. I carry immense gratitude and love for her, and I dedicate this work to her memory. I am also deeply grateful to my wonderful wife, Snehal Bhosale, for her patience and understanding in taking such loving care of our son, Shreejal Bhosale, allowing me the time and focus needed to complete this work. I thank my little boy for his endless joy and innocence, which have been a constant source of motivation and inspiration throughout this journey.



# Abstract

Industrial control systems (ICSs) are a critical infrastructure and have continuously evolved, replacing the physical control mechanism with automated systems. Previously, ICSs were more prone to safety incidents; recent advancement has made ICSs more prone to security incidents. Risk management has been a primary focus for safe and secure functioning in industrial operations. Currently, risk management methods treat the process separately for safety and security and depend on manual assessments, which are slow and can cause unintentional errors, leading to a wrong assessment.

To address the challenges of fragmented and manual safety and security risk assessments in ICSs, this thesis proposes an integrated, (semi-)automated methodology tailored to the unique demands of modern industrial automation environments. Recognizing the increasing complexity introduced by Industry 4.0, Industrial IoT, and digital transformation, the approach unifies safety and security perspectives using a Bayesian Belief Network (BBN)-based framework. It enables probabilistic modeling of uncertainties and interdependencies within ICS components, crucial for anticipating cascading failures or security breaches. The methodology is further strengthened by advanced information modeling techniques, including AutomationML (AML), Asset Administration Shells (AAS), and semantic ontologies, which create a structured, interoperable "single source of truth" for risk-relevant data.

A practical implementation is demonstrated through a modular production system use case, integrating data from system architecture, stakeholder analysis, and vulnerability scans. Python serves as the backbone for automating key processes such as: data extraction, model generation, and probability inference, achieving a significant level of automation while ensuring adaptability to heterogeneous ICS environments. The research contributes a flexible framework that aligns with ISO 31000 standards, provides actionable insights for stakeholders, and lays the groundwork for future advancements, such as machine learning integration, real-time monitoring, and predictive maintenance via digital twins.

**Keywords:** industrial control systems, risk assessment, Bayesian belief network, AutomationML.



# Contents

<b>Abstract</b>	<b>ix</b>
<b>Contents</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Motivation . . . . .	4
1.2 Problem Statement . . . . .	5
1.3 Research Objectives . . . . .	5
1.4 Contributions . . . . .	7
1.5 Further Chapters and Structure . . . . .	8
<b>2 Background and Literature Survey</b>	<b>11</b>
2.1 Risk Identification . . . . .	12
2.2 Risk Assessment . . . . .	14
2.3 Risk Treatment . . . . .	19
2.4 Risk Management . . . . .	20
2.5 Information Model for Risk Assessment-Single Source of Truth . . . . .	21
2.6 Human Errors in Risk Management . . . . .	23
2.7 Automation process and levels . . . . .	23
<b>3 Safety, security, and their relationship</b>	<b>25</b>
3.1 Safety . . . . .	25
3.2 Security . . . . .	26
3.3 Similarities and differences between safety and security . . . . .	27
3.4 Relationship . . . . .	28
3.5 Benefits of integrating safety and security . . . . .	30
<b>4 Methodology</b>	<b>31</b>
4.1 Problem identification and motivation . . . . .	31
4.2 Objectives of the solution . . . . .	32
4.3 Design and development . . . . .	32
4.4 Demonstration . . . . .	33
4.5 Evaluation . . . . .	35
4.6 Communication . . . . .	35
	xi

<b>5</b>	<b>Information Collection Phase</b>	<b>37</b>
5.1	Data Input Sources . . . . .	37
5.2	Automation possibility of sources . . . . .	42
5.3	Limitation or challenges . . . . .	42
<b>6</b>	<b>Information Organization Phase</b>	<b>45</b>
6.1	Automation Markup Language . . . . .	45
6.2	Asset Administration Shell . . . . .	58
6.3	Ontology . . . . .	63
6.4	Information Utilization . . . . .	68
<b>7</b>	<b>Risk assessment using Bayesian Belief Networks</b>	<b>71</b>
7.1	Single source of truth Model . . . . .	71
7.2	Introduction to BBN . . . . .	73
7.3	Automation Markup Language (AML) and Bayesian Belief Network (BBN) integration using Python . . . . .	75
7.4	BBNs implementation . . . . .	76
7.5	Node Probabilities Probability of Node (P(N)) and Severity Severity of Node (S(N)) . . . . .	81
7.6	Individual Node Risk Risk of Node (R(N)) . . . . .	83
7.7	Relation between nodes . . . . .	84
7.8	Conditional Probability Table (CPT) Assignment Algorithm for Child Node . . . . .	86
7.9	Program Flow and Representation . . . . .	90
<b>8</b>	<b>Risk Mitigation Strategy using Bayesian Belief Network</b>	<b>93</b>
8.1	Sources for treatment measures . . . . .	93
8.2	Risk Treatment Procedure . . . . .	99
8.3	Identified mitigation for the risk . . . . .	101
8.4	Integration of Mitigation node in BBN implementation . . . . .	103
8.5	Mitigation probabilistic relationship . . . . .	104
8.6	Representing BBN Graphs . . . . .	106
<b>9</b>	<b>Results and Discussion</b>	<b>109</b>
9.1	Risk assessment results . . . . .	109
9.2	Evaluation based on standard ISO 31000 . . . . .	114
9.3	Evaluation based on automation capability . . . . .	116
9.4	Advantages of the Integrated Risk Assessment Method . . . . .	118
<b>10</b>	<b>Conclusion and Future Work</b>	<b>121</b>
	<b>List of Abbreviations</b>	<b>127</b>
	<b>List of Figures</b>	<b>129</b>

<b>List of Tables</b>	<b>131</b>
<b>List of Algorithms</b>	<b>133</b>
<b>Bibliography</b>	<b>135</b>



# Introduction

The term Industrial Control System (ICS) is used collectively for Industrial Automation and Control System (IACS) (coined by the International Society of Automation (ISA), 2002), Supervisory Control and Data Acquisition (SCADA), Process Control Systems (PCS), and Distributed Control Systems (DCS) [125]. Depending on the application and industry, each ICS functions differently and works to manage control tasks efficiently [27]. Traditional ICSs consist of various components with hardware and software for directing, controlling, regulating, and managing (critical) industrial processes, including manufacturing, product handling, production, and distribution. Their goal is to execute open and closed-loop control with associated instrumentation and industrial communication systems with various protocols [58]. Hierarchically, the technologies and devices range from supervisory control for geographically distributed networks such as SCADA, Human Machine Interfaces (HMIs), and data acquisition systems (Data Historian) that receive data from control level devices such as Programmable Logic Controllers (PLCs) and Remote Terminal Units (RTUs) and field level devices such as sensors and actuators [122]. ICSs monitor and control equipment used in different industries such as the power sector, nuclear energy, oil and natural gas, and automotive. An ICS is also part of the Critical Infrastructure (CI), and its failure can lead to adverse effects on the safety of people, processes, and environments and can leave the company vulnerable to criticism [84].

Operation of ICSs consist of cooperation between humans, hardware, and software to perform the necessary functionality. Figure 1.1 represents the hardware, software, and human assets that influence the functionality of the plant in fulfilling the goal. The process is influenced by Enterprise Resource Planning (ERP) and Manufacturing Execution System (MES) in making decisions related to business processes, product planning, or plant maintenance. The hardware and software parts of ICSs are interlinked in some cases, and failure of one can lead to the failure of the other. In recent times, some of the hardware on location, such as data storage devices (Historian), SCADA, has

been updated to a cloud-based operation. This makes it easier for remote communication, control, and data collection.

The main blocks of the metamodel in Fig. 1.1 consist of primary requirements of the industry with a focus on safety and security risk assessment, the typical human position involved, and hardware in combination with software and communication modules. The blocks are defined as follows:

- **User:** User is the human element that is necessary for overall safe and secure operation. The Industry 4.0 (I4.0) aims to automate the complete operation of the industry, but the *User* remains a necessity and a difficult part to automate. *User* can be *internal* person (e.g. company employee) or *external* person (e.g consultant, service provider). They are a vital part of the process and use tools of *Hardware* and *Software* to carry out the necessary function based on the requirements. *User* is an experienced or trained person responsible for building, operating, and maintaining an ICS. It includes *Developer, Operator, Risk manager, Worker or Labourer, Maintenance Admin, and others*.
- **Hardware:** Hardware refers to the tangible asset necessary for the process. It includes the field devices, industrial equipment (*equipment that is a part of the production, manufacturing, etc. process, but is not the produced material itself. It goes up to the controller layer of the automation pyramid. It includes process equipment (conveyer belts, boilers tanks), field devices (sensors, actuators), cobots, robots, machines (milling, drilling), controllers (PLC) etc.*) HMI panels, work stations or laptops, networking devices (switches, routers), etc. Hardware makes it possible for humans to control the process with touch panels and emergency buttons. Hardware is a tool in the hands of *User* to communicate using software to all the stations, and also to collect information and visualize it at a single station.
- **Software:** Software is a collective term used to define the non-tangible assets of a system. It controls the hardware and complete operation of instruments to achieve a specific purpose. E.g., SCADA software, firmware, OT adapter, data historian, and process logic. Software is very crucial with regard to security, as many security incidents try to affect the software, and thus, normal operation is hampered.
- **Communication:** Communication includes the protocols, interfaces, and channels used for data exchange between various components in the system, including between hardware, software, and users. It involves both physical transmission and logical layers. Secure and reliable communication is vital for ensuring timely response, remote accessibility, and data integrity in real-time industrial operations. Communication acts as the backbone that interconnects all other elements and is a frequent target of cyber threats.
- **Requirements:** Requirements define the functional and non-functional needs that the system must fulfill to ensure safe and secure operation. To ensure that the

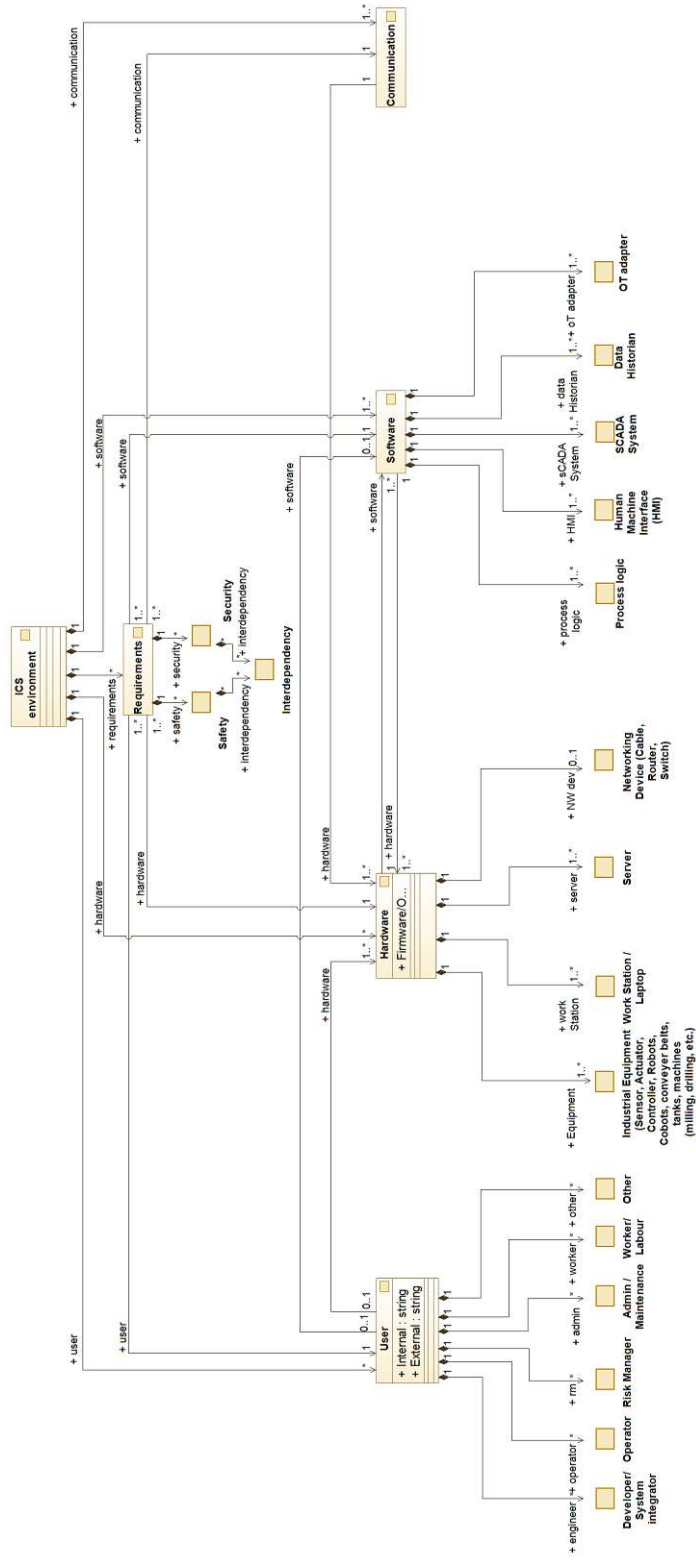


Figure 1.1: ICS general structure

industrial system operates without causing unacceptable risk of physical harm or damage, safety requirements are implemented. They include fail-safes, redundancy, physical access control, emergency shutdown systems, and compliance with safety standards like IEC 61508 or ISO 13849. In order to focus on protecting the system from intentional malicious attacks, unauthorized access, data breaches, and cyber-physical threats, security requirements are necessary. They include authentication, authorization, encryption, secure boot, vulnerability management, and compliance with security standards like IEC 62443. A breach in security can compromise safety, and unsafe configurations may be exploited as security vulnerabilities. Therefore, interdependency between the safety and security requirements needs to be considered during the design and implementation phases.

### 1.1 Motivation

Information Technology (IT) focuses on protecting data, while Operational Technology (OT) ensures the continuous operation of processes. The difference results in varying definitions of critical functions and risk priorities [78]. IT and OT convergence, first discussed in the 1990s and increasingly relevant with I4.0 and the Industrial Internet of Things (IIoT), aims to integrate these systems for improved efficiency and data communication. Despite some similarities in technology, such as PC-based hardware and IP-based communication, key differences remain in priorities, architecture, and requirements [122]. As these systems integrate, ICSs face exposure to cyberattacks and security breaches. A shift in focus is necessary from safety against accidental failures to addressing cybersecurity vulnerabilities, which can directly compromise safety [133]. According to NIST 800-82 [122], these differences span areas such as timeliness, performance, availability, risk management, physical effects, system operation, and component lifetimes.

Recent data underscores the growing security risks to ICS. A Kaspersky report [1] indicates that approximately one-third of ICSs were attacked, primarily via the internet, often with malicious intent. These attacks are increasing due to geopolitical motives, financial incentives, and the convergence of IT and OT systems [31]. Notably, the US gas pipeline attack revealed the critical vulnerabilities of these systems. In response, the World Economic Forum recommended implementing holistic risk management strategies that integrate safety and security to monitor, respond to, and prevent such incidents [103].

Traditionally, risk assessment, comprising risk identification, analysis, and evaluation, has been conducted manually and separately for safety and security. However, the growing interdependence between safety and security demands an integrated approach. Such a model can capture the complex relationships among threats, vulnerabilities, hazards, and component failures [138]. The manual nature of risk assessment introduces significant limitations. It is time-consuming, error-prone, and lacks repeatability and reliability, especially when managing complex systems. Automating the risk assessment process is

essential to overcome these limitations. Automation can reduce dependence on manual processes, improve consistency, enhance performance, and enable better decision-making [96].

An integrated and (semi-)automated approach to safety and security risk assessment not only addresses the challenges posed by IT/OT convergence but also ensures a robust, efficient, and scalable mechanism for protecting critical infrastructure.

## 1.2 Problem Statement

The risk assessment process for safety and security in ICSs have traditionally been conducted separately, as these domains have evolved independently. While risk assessment methods for process and functional safety are well-established, they often fail to account for the growing influence of security threats introduced by engineering advancements. This disconnect has necessitated new regulations, such as the Network and Information Security Directive (NIS2), Cyber Resilience Act (CRA), and the updated Machine Regulation (2006, revised in 2023), as well as integrated standards like IEC 63069 and 63071, to address the convergence of safety and security concerns.

Despite these efforts, several engineering challenges remain in achieving integrated risk assessment. These include identifying and integrating diverse data sources for safety and security, retrieving relevant and actionable information, aligning disparate safety and security frameworks, and developing unified methods for risk assessment and management. Additionally, the dynamic nature of cyber threats requires adaptive and scalable solutions that current approaches struggle to deliver.

Manual risk assessment methods further exacerbate these issues, as they are time-consuming, error-prone, and unable to effectively capture interdependencies between risks, where one risk can propagate or amplify others. As systems become increasingly complex, the limitations of manual approaches hinder the identification and mitigation of relevant risks. The integration of safety and security considerations in a unified risk assessment remains critical yet underexplored.

Automating the risk assessment process offers a potential solution to these challenges by enhancing efficiency, consistency, and the ability to manage complex interdependencies. However, the accuracy, reliability, and trustworthiness of automated systems remain critical concerns, requiring rigorous validation mechanisms. Addressing these gaps is essential to ensure robust, unified safety and security risk management, enabling the protection of ICSs against emerging cyber-physical threats.

## 1.3 Research Objectives

The research aims to establish a structured approach divided into conceptual frameworks and proof-of-concept implementations to achieve specific objectives efficiently. The goals include conducting a state-of-the-art survey to identify data sources relevant to

security-related data collection and developing a prototype for continuous data collection. Additionally, the research seeks to design a concept for (semi-)automated risk identification and dynamic risk assessment models, followed by the creation of a prototype for their implementation. Furthermore, it aims to develop a model for generating measures that incorporate methods for automating risk treatment. Finally, the research intends to validate this model and assess its limitations.

The research goal is summarized using the research question,

*What mechanisms can improve the process of risk management and conduct it in a (semi-)automated way to guarantee ICSs' safe and secure operation?*

The research question is further refined to understand the contributions.

1. What are the data input sources for safety and security-relevant data applicable in ICSs environment?  
→ It identifies the data sources that are available in the ICSs environment and are necessary for safety risk assessment and security risk assessment. It includes operational data, logs, environmental data, and contextual information. The outcome ensures that relevant and comprehensive data inputs are available for subsequent risk assessment and management processes.
2. What are the safety and security risk indicators based on the collected data?  
→ It identifies key risk indicators derived from the collected data, establishing measurable metrics that signal potential risks. The indicators bridge the gap between raw data and actionable insights, forming a foundation for predictive and dynamic risk assessment in ICSs environments.
3. Which information models are most suitable for supporting the automation of risk assessment and management in ICSs?  
→ It helps to select robust information models that enable efficient storage, retrieval, and processing of safety and security-related data. The models facilitate automation by structuring data in a way that supports risk identification, assessment, and management tasks, ensuring scalability and adaptability in dynamic ICSs contexts.
4. Which methods are suitable for an integrated safety and security risk analysis and assessment?  
→ It includes a comprehensive evaluation of existing and novel methods for integrated risk analysis, considering the interdependence between safety and security risks. It leads to developing or refining techniques that holistically assess risks, accounting for operational disruptions and security breaches in ICSs.
5. What methodology can be used to identify and implement risk treatment strategies?  
→ It proposes a methodology that enables systematic identification and implementation of risk treatment strategies. The strategies include preventive measures, mitigation plans, and recovery protocols tailored to the unique requirements of ICSs environments. The methodology also incorporates automation to enhance efficiency and rapidly respond to evolving risks.

Addressing the research questions provides a comprehensive framework for advancing risk assessment and management automation in ICSs. It contributes to developing data-driven insights, integrated risk analysis methods, and actionable risk treatment strategies, ensuring safer and more secure ICSs operations.

## 1.4 Contributions

In our opinion, this research makes significant contributions in the field of risk assessment in ICSs. The research methodology contributes to the systematic identification, assessment, and treatment of both safety and security risks in an integrated framework. By leveraging conceptual models, automated processes, and data-driven insights, the methodology ensures a comprehensive understanding of risks while reducing redundancies between safety and security domains. The key contributions of this research are as follows:

- **Stakeholder Analysis to Understand the Needs of the Industry:** Given the complexity of ICSs, it was essential to understand the distinct safety and security priorities of each group. Through a structured stakeholder analysis, the research was able to identify and classify these priorities, forming a clear picture of their expectations and concerns. This foundational understanding directly informed the development of a safety and security methodology tailored to stakeholder requirements and aligned with prevailing regulatory frameworks.
- **Identification of Data Sources for Safety and Security Information:** Compiled and categorized relevant data sources critical for safety and security in ICSs. We have identified the relevant sources and data we need for the assessment, where it is stored, the frequency at which the data is needed and retrieved, and the automation presence in the storage and retrieval process [15]. The data can be system-related information (*assets and connections*), asset-related information (*Mean Time Between Failure (MTBF), failure rate*), process definition (*normal operation of the system*), and threat intelligence (*vulnerabilities, their probabilities*). The systematic identification ensures comprehensive coverage of both safety and security aspects.
- **Data Integration Through Information Models:** Utilized industry-standard information models, including AML and the Asset Administration Shell (AAS), to structure and represent data for risk assessment effectively. These models ensure interoperability and support the automation of data-driven risk management processes [14, 16]. An analysis that explores information models, comparing AAS and AML, to evaluate their strengths, weaknesses, and applicability in enhancing risk assessment accuracy, efficiency, and automation [20].
- **Establishing Relationships between Safety and Security Risk Assessment Elements:** Developed an ontology-based framework to analyze and model the

interdependencies between safety and security in ICS. This ontology serves as a structured and formalized representation of the relationships between safety and security risk elements, enabling a unified understanding of their interactions. The framework highlights how safety measures can impact security outcomes and vice versa, capturing both synergistic and conflicting dynamics. The ontology serves as a unified source of truth, facilitating the integration, organization, and retrieval of risk-related information [17].

- **Integrated Risk Assessment Framework:** Developed a novel methodology that combines safety and security risk assessment into a unified framework using BBN [14, 18, 19]. This approach enables a holistic analysis of interdependent risks, improving the reliability and robustness of assessments in ICSs.
- **Automation of Risk Assessment:** Designed and proposed an automated risk management system to address the inefficiencies and limitations of manual risk assessment methods. This system integrates diverse data sources, enabling seamless data collection, processing, and analysis to support real-time and scalable risk management [19].
- **Concept and Implementation of Risk Treatment:** Developed a structured methodology for identifying, prioritizing, and implementing risk treatment strategies in ICSs. A strategy that highlights automating risk treatment and its integration using models is developed [24]. This includes implementing preventive, mitigative, and recovery strategies, ensuring adaptability to evolving risks while maintaining operational continuity.

### 1.5 Further Chapters and Structure

The thesis is organized into several interconnected chapters, each addressing a critical aspect of the integrated safety and security risk assessment methodology. The structure is designed to lead the reader through the background, methodology, implementation, and evaluation stages systematically.

The thesis begins with a comprehensive literature survey in Chapter 2. It reviews existing approaches in safety and security risk assessments, particularly focusing on their integration. It highlights the current state-of-the-art techniques, identifies research gaps, and provides the rationale for developing a unified framework. It also focuses on the international standards that are applicable for the evaluation phase.

Following this, the integration and interaction strategies in Chapter 3 explore how safety and security considerations can be combined in a coherent manner. It delves into the conceptual and practical aspects of integration, discusses possible synergies, and analyzes potential conflicts. The benefits and limitations of these strategies are also critically examined.

Next, the methodology in Chapter 4 outlines the proposed approach for implementing and evaluating the integrated risk assessment framework. It includes an overview of the workflow and its components, emphasizing applicability in industrial contexts and adaptability to different domains.

The information gathering phase introduced in Chapter 5 highlights the process of collecting and organizing relevant data for risk assessment. It symbolises a single source of truth for the information needed for the risk assessment. The chapter discusses various information modeling standards and frameworks, such as AML, AAS, and Ontologies. It also presents the development of a customized information model to support integrated safety and security analysis, serving as a foundation for subsequent phases.

Building on this, the information extraction phase in Chapter 6 explains how relevant data is retrieved from the developed information model to be used in the integrated risk assessment. The step ensures that both safety and security aspects are adequately represented and analyzed. The information is then organised for its use in integrated risk assessment.

The risk assessment methodology in Chapter 7 details the use of BBN for conducting the actual assessment. It justifies the use of BBNs as a probabilistic reasoning tool capable of handling uncertainty and interdependencies between various factors in safety and security domains.

Subsequently, the risk treatment phase in Chapter 8 continues with the application of Bayesian reasoning to support decision-making. Based on the assessment outcomes, this chapter explores how risks can be mitigated through informed strategies that consider both safety and security implications.

The thesis then presents the results, evaluation, and discussion in Chapter 9. It showcases the findings obtained from applying the proposed methodology. The effectiveness, limitations, and potential for improvement of the approach are critically analyzed here.

Finally, the conclusion in Chapter 10 summarizes the key contributions of the research, reflects on the insights gained, and outlines directions for future work in the domain of integrated safety and security risk assessment.



# Background and Literature Survey

A general risk management concept is a part of many fields, such as finance, safety engineering, health monitoring, enterprise, transportation, security, and supply chain management. For ICSs, various standards, frameworks, and best practices are available for the safety or security needs of the industry. For instance, IEC 62443 [55] and NIST SP 800-82 [122] specifically target only security aspects of an ICS, while IEC 61508 [54] and ISO 12100 [59] address only safety.

The standards deal with the safety or security requirements on a broadly generic level. Although periodically conducting a risk assessment is a best practice, performing it manually is time-consuming and can often lead to errors. Also, risk assessment in the realm of ICSs was mainly related to safety. There is still development in the IEC community that would serve a combination of safety and security risk assessment. IEC TR 63069 (technical report) explains and guides the typical applications of IEC 61508 and IEC 62443 in industrial process measurement, control, and automation [56]. IEC TR 63074:2019 will guide the use of IEC 62443 security-relevant threats and vulnerabilities that influence the functional safety of the control system [57].

Frameworks like MITRE ATT&CK<sup>1</sup> and D3FEND<sup>2</sup> provide strategies for understanding and mitigating attacks. Defining some important terms is necessary to understand the importance of risk management and the current state-of-the-art. According to [61], risk and related terms are defined as follows.

- Risk: Uncertain event hampers objectives (goals or outcomes).
- Risk identification: Process of recognizing and describing risks.
- Risk analysis: Process of determining the nature and level of risk.

---

<sup>1</sup><https://attack.mitre.org/>

<sup>2</sup><https://d3fend.mitre.org/>

- Risk evaluation: Process of comparing risk analysis results with risk criteria to determine acceptable or tolerable levels of risks.
- Residual risk: The risk that remains after risk treatment.
- Acceptable risk: Informed decision of the organization to go ahead with the risk, based on analysis.
- Tolerable risk: Organization's readiness to bear risk after the risk treatment to achieve its objectives.
- Risk assessment: Identification, analysis, and risk evaluation process.
- Risk management: Continuous process of providing risk assessment based on risk reduction options.

### 2.1 Risk Identification

The first step of risk management requires understanding the System Under Consideration (SUC). The data needed for the beginning of the process is critical and is obtained mainly qualitatively. It includes system architecture, asset inventory, network segmentation, Overall Equipment Efficiency (OEE), performance report, use of the component, previous analysis report, etc. [132]. According to [55], historical security incidents and other information sources are used for threat identification. Prior vulnerability assessments and databases are used as the source of information for vulnerability knowledge.

According to [35], the risk identification phase should answer questions like: *who/what is the risk agent?* (e.g., attackers, failure, workers), *why is the agent motivated?* (Agent's intent, unintentional mishap, component efficiency), *what is at risk?* (e.g., assets, people, process, environment), *how will the attack take place?* (e.g., scanning, attacker capability), *where is the component located?*, and *when is the component going to fail?* (e.g., exploited vulnerabilities).

The standard methods in risk identification within the organization include but are not limited to brainstorming, documentation review, business impact review, assumption analysis, Delphi technique, root cause analysis, SWOT (Strengths, Weaknesses, Opportunities, and Threats) analysis, expert judgment, etc. [130]. The risk identification techniques mentioned in [109] are categorized according to safety and security. They consider intentional misuse, unintentional mishap, hazard, attack, threat, vulnerability identification, or safety and security requirements, but exclude frameworks or combinations of methods to evaluate risks and their probability. Table 2.1 highlights the methods mentioned in [109] and what questions are referred to in the output [35].

The risk identification information gathering is qualitative based on the input requirement and, hence, difficult to automate. However, efforts are made to automate security-based risk identification, while safety-based risk identification remains the same. In [40, 15], the

Method	Input	Focus	Answers
Functional Hazard Analysis	functional flow diagram, design and operational knowledge	Safety	Who, What, Where, When
Preliminary Hazard Analysis	design, hazard knowledge, preliminary hazard list, top-level mishaps	Safety	Who, Why, Where
Hazard and Operability Methodology (HAZOP) [110]	design models	Safety	Who, What, Where
Failure Modes and Effects Analysis (FMEA) [49]	system design, operational constraints, success and failure boundaries, credible failure modes and probability	Safety	Who, Why, What, Where
Fault Tree Analysis (FTA) [113]	drawings, schematics, procedures, functional block diagrams	Safety	Who, Why, What, Where
Keep All Objects Satisfied Security extension (KAOS SE)	stakeholder interviews and documents, system model	Security	Who, Where, What, How, When
Misuse Cases	functional requirements (use case for input)	Security	Who, What, How
Attack Trees (AT) [128]	attacker's Skill, scenarios, motivation	Security	Who, What, Where, How, Why, When.
Threat Trees	complete set of threats, system description	Security	What, When

Table 2.1: Risk identification methods [109], [35]

author mentions the challenges of automating the alignment of safety and security risk identification and assessment throughout the lifecycle of a modular production system. In [38], the author has proposed a method to automatically identify only security risks and potential consequences to automatically generate an attack graph using engineering data in a Cyber Physical System (CPS).

In [14], a risk identification and overall risk assessment method is proposed using AAS. Various mathematical models (e.g., Bayesian-based, steady state analysis), algorithms (e.g., SVM, decision trees, text mining, deep learning), modeling techniques (e.g., automated Hazard and Operability Methodology (HAZOP), simulations), and models (e.g., Petri Nets, UML) have been used in recent times for proactive risk identification. Behavioral Anomaly Detection (BAD) is a method recommended to identify existing risks. [86] mentions following BAD methods:

- Network-Based Behavioral Anomaly Detection,

- Agent-Based Behavioral Anomaly Detection, and
- Historian-Based and Sensor-Based Behavioral Anomaly Detection.

### 2.2 Risk Assessment

Risk identification is followed by risk analysis. Some methods mentioned in Table 2.1 such as Fault Tree Analysis (FTA), Attack Tree Analysis (ATA), Failure Modes and Effects Analysis (FMEA), HAZOP are further extended to obtain the complete assessment. [132] reviewed different risk analysis methods used in plants and highlighted their types of inputs ((a) *Plans or diagrams*, (b) *Process and reactions*, (c) *Asset function and quantity*, (d) *Probability and frequency*, (e) *Implemented policy*, (f) *Environment*, (g) *Documentation, and Historical data*). Moreover, the following types of standard methodology are named in the publication (a) *Deterministic*, (b) *Probabilistic*, (c) *Qualitative*, and (d) *Quantitative*. Finally, 62 techniques are categorized based on those methodologies, and expected output ((a) *Management actions and recommendation*, (b) *List of failures, hazards, threat, and vulnerability*, (c) *Probabilistic failure rate, accident frequency*, (d) *Hierarchisation like severity, criticality or type of risk*).

In [93], the author has evaluated the existing methods for risk analysis in ICSs. The methods such as CORAS, OCTAVE, HAZOP mentioned in the report were primarily qualitative, and a new qualitative framework was proposed called Modified Risk Analysis Framework for ICSs (MRAF-ICS). It uses a threat modeling approach for threat and vulnerability identification. Risk analysis can be qualitative, quantitative, or both, as mentioned in [132], and finding suitable methods from Table 2.2 is a vital step in analysis and assessment.

Risk assessment is performed with some of the methods mentioned in Table 2.2 [138]. The methods are divided into three categories: safety-based, security-based, and integrated risk assessment techniques based on their primary target in the industry. The methods are further categorized as static (S), dynamic (D), quantitative (Qn), or qualitative (Ql) methods to understand them better and to choose wisely for the application. The impact and use in current threat scenarios are some of the comparison points of the methods, and one needs to consider them before choosing the method.

#### 2.2.1 Safety Risk Assessment

Safety is a core principle guiding policies, regulations, and management across energy, chemicals, transportation, water, and health sectors. Organizations have become increasingly focused on safety due to the rise of hazardous technologies and activities. Research links hazards to evolving organizational features, suggesting that major accidents may be inevitable in high-risk systems [102]. The perspective highlights the limits of safety and the role of organizational competency. According to [2], the human factor remains the weakest link in ICS environment.

Reference	Assessment method	Type *	Focus
[113]	Fault Tree Analysis (FTA)	S, Ql, Qn	Safety
[49]	Failure Modes and Effects Analysis (FMEA)	S, Qn, Ql	Safety
[80]	Systematic Theoretic Process Analysis (STPA)	S, Ql	Safety
[110]	Hazard and Operability Methodology (HAZOP)	S, Ql	Safety
[89]	Goal Tree-success Tree and Master Logic Diagram (GTST-MLD)	S, Qn	Safety
[128]	Attack Tree Analysis (ATA)	S, Qn, Ql	Security
[141]	Systematic Theoretic Process Analysis Security (STPA-sec)	D, Ql	Security
[53], [144]	Bayesian Network approach	S, Qn, Ql	Security
[78], [77], [79]	Boolean logic-driven Markov Processes (BDMP)	D, Qn, Ql	Combined
[29]	Non-functional Requirements (NFRs)	S, Ql	Combined
[43]	Systematic Theoretic Process Analysis - Safety and Security (STPA-SafeSec)	S, Ql	Combined
[105]	Six-step Model (SSM) and Information Flow Diagram (IFD) integration approach	S, Qn, Ql	Combined
[72], [45], [144], [146]	Bayesian Belief Network (BBN)	D, Qn, Ql	Combined

\* [S: Static, D: Dynamic, Qn: Quantitative, and Ql: Qualitative analysis]

Table 2.2: Risk analysis and assessment methods

Various methods and tools are developed to address safety in ICS. FTA, a qualitative and quantitative failure mode impact analysis, is used to model complex failure modes of the system concerning component failures. The values (reliability, safety) are calculated based on the system information. The drawback of this method is its implementation cost and difficulty in finding all possible accidental paths [113]. FTA is a top-down and deductive method, while FMEA is a bottom-up and inductive approach. The time required for failure analysis with this method is considerably higher and incomplete, as it does not consider the relationship between different failure components. It is limited to analyzing a single cause of impact [49].

HAZOP helps analyze and review industrial processes to detect any deviation from the design conditions. It relies heavily on expert opinion and fails to highlight the issues. This method cannot handle the numerous and complex interactions that occur with I4.0 [110]. Goal Tree Success Tree - Master Logic Diagram (GTST-MLD), a goal-oriented method, implements prior knowledge to identify hazards. It represents all the elements of the

system hierarchically. GTST-MLD ignores undesired interactions and cannot adequately analyze vulnerabilities, failures, and errors [89, 138].

Apart from assessment methods, ontological models and tools have also been developed to perform safety risk assessments. An ontology-based fault diagnosis method for industrial applications is detailed in [87]. To identify hazards and their causes, [12] proposes an ontology-based approach with a generic model that translates adverse process states into visual pathways. A runtime safety management model based on OSHA, featuring the RAMIRES tool (Risk-Adaptive Management in Resilient Environments with Security), is introduced in [127].

### 2.2.2 Security Risk Assessment

Safety and security share several commonalities and differences, which are crucial when developing policies and measures. Their primary similarity lies in the mutual goal of risk prevention and ensuring the seamless operation of ICS. As noted in [78], these fields overlap in aspects such as risk management, imposing constraints, implementing safeguards, and defining requirements.

ATA, an independent module, is widely used for security assessments. It presents the steps for an attack in a graphical manner. There are no standards for attack tree construction. ATA depends on known attacks from historical data and intuitive expert opinion for attacks. System Theoretic Process Analysis (STPA) is a technique used for hazard analysis based on System Theoretic Accident Model and Processes (STAMP). STPA aims to identify the causes of hazards and accident scenarios along with the entire accident process [80]. STPA-sec is an extension of STPA for security issues [141].

Similar to safety approaches, the world of security has also used ontological approaches to provide the relation between different security components. An ontology-based approach to information security is proposed in [9], advocating using software agents and semantic knowledge representation for cost-efficient security management. In [92], an ontology-based cybersecurity framework is introduced to enhance IoT security by identifying threats through knowledge reasoning and proposing suitable countermeasures.

Ontology-driven cyber-attack detection for SCADA systems is discussed in [75], though it relies on a fundamental alert-based ontology requiring further development. To address resource constraints and evolving threats, [126] suggests leveraging ontology-based knowledge for security risk assessments, focusing on ICS security knowledge and its lifecycle.

### 2.2.3 Integrated Safety and Security Assessment

Integrating safety and security is vital for risk assessment in critical systems. Surveys [78, 136, 32, 28, 120] have explored integrated safety and security models, highlighting their applicability in real-world scenarios. Ontology-based methodologies [7, 101] have

been proposed for hybrid RAMSS (Reliability, Availability, Maintainability, Safety, and Security) risk assessment.

Combined system-theoretic approaches, such as STPA and STPA-Sec, are used to form STPA-SafeSec and address safety and security concerns. STPA-SafeSec, as the name suggests, treats safety and security equally and performs integrated analysis for optimum results. A single framework in STPA-SafeSec considers the factors influencing safety and security issues. Although STPA-SafeSec can be used to analyze the interdependencies of safety and security, quantification of risks is not possible [43] [138].

Frameworks like the V-model [73] and model-based approaches [76, 82] further enhance this integration. Boolean logic-driven Markov Processes (BDMP) performs quantitative and qualitative analysis for safety and security risk assessment. BDMP are dynamic processing models. The actual situation and network structure assumptions significantly impact the results. Some wrong assumptions about network elements will make the analysis inaccurate [78]. In [77], the author has presented risk evaluation using BDMP, combining safety and security. They have also compared the benefits of an integrated approach and characterized interdependencies between safety and security using a use case of the industrial pipeline.

Non-functional Requirements (NFR), a qualitative evaluation method, answers the questions related to fulfilling safety and security requirements. NFR must be repeated after every system update, which cannot be economical [29]. BBN is a probabilistic graphical model that utilizes Bayesian inference for conditional dependency [144]. BBN [144, 146] enable dynamic and probabilistic risk assessments. Bayesian inference helps to assess the risk under unknown threats and hazards. The apparent challenges in building BBN models are a lack of historical data and prior knowledge of a large number of attacks [72] [138]. BBN [144, 146, 74] play a key role in assessing safety and security in complex systems. For instance, [74] presented a case of Bayesian-based assessment on CPS.

Various methodologies facilitate the integration of safety and security. Ontology-based approaches [7, 101, 101] use semantic frameworks to identify and mitigate risks. Standards such as IEC TR 63069a and 63074 [67] bridge functional safety and cybersecurity. Techniques like Security Vulnerability Analysis (SVA) and Layer of Protection Analysis (LOPA) [50] align safety with cybersecurity. Dynamic event trees [30] and consequence-based approaches [26] address risks arising from intentional and accidental factors. Hybrid models, including RAMSS ontologies [7], enhance the reliability of integrated assessments.

The primary challenge of risk assessment is detecting multiple component failures and evaluating the result. This problem can be solved using an BBN as it graphically detects the problem and computes probability using interdependency [146]. Although the methods mentioned in Table 2.2 contain combined methods for safety and security, very little or no attempt is made to use those methods in a combined fashion.

Other techniques that have been used for risk assessment include the Analytic Hierarchical Process (AHP) [111], Monte Carlo Simulation [137], etc. Monte Carlo simulation, a powerful tool, is used for evaluating risks. It can analyze the adverse effects and provide

an estimation with confidence interval calculations based on data or expert opinion. Just like in an attack tree, the user needs to assess all scenarios, which is complicated. Although it makes complex calculations with multiple simulations, it fails to consider multiple component failures and hence cannot represent the dependency properly.

Despite progress, integrating safety and security presents challenges. Issues such as data reliability and model complexity [15] hinder implementation. The dependencies between safety and security increase both costs and complexity [106]. Surveys [104] emphasize the need for robust frameworks to overcome these barriers.

The comparison of the methods is performed on the following basis using [70] [43] [85] [114], [6], [83], [112]:

1. Domain (a) only safety, (b) only security, (c) integrated
2. (a) probabilistic, or (b) deterministic
3. methodology (a) qualitative, and (b) quantitative,
4. nature of analysis (a) static, and (b) dynamic,
5. addressing unknown threats and hazards,
6. possibility of Model-Based analysis
7. Interdependencies (\*! defines assessment only security to safety is possible) (a) Antagonism, (b) Conditional dependency, (c) Mutual reinforcement, and (d) Independence or (O) Addresses no interdependency
8. Complexity handling
9. Intelligent implementation
10. Automation possibility

Table 2.3: Risk assessment methods

Methods ↓ Basis →	1	2	3	4	5	6	7	8	9	10
FHA	a	b	a	a	N	N	O	N	N	N
PHA	a	b	a	a	N	N	O	N	N	N
FTA	a	a	a, b	a	N	P	O	P	P	P
FMEA	a	b	a, b	a	N	P	O	P	P	P
HAZOP	a	b	a	a	N	N	O	N	N	N
STPA	a	b	a	a	N	N	O	N	N	N
STPA-sec	b	b	a	b	N	N	O	N	N	N
STPA-safesec	a, b	b	a	b	N	N	*!, a, b, c, d	P	N	N

–Continued on next page

Table 2.3 – continued from previous page

Methods ↓ Basis →	1	2	3	4	5	6	7	8	9	10
ATA	b	a	a, b	b	N	P	O	P	P	P
OCTAVE	b	b	a	b	N	N	O	N	N	N
CORAS	b	b	a	b	N	N	O	N	N	N
FMVEA	c	a	a, b	b	N	N	a, b, c, d	N	P	N
Bow Tie Analysis (BTA)	a, b, c	a	a, b	a	N	N	*!, a, b, d	N	P	N
BDMP	a, b, c	a	a, b	b	N	P	a, b, c, d	P	P	P
BBN	a, b, c	a	a, b	b	P	P	a, b, c, d	P	P	P

\*[P- Possible, N - Not enough information, O- Addresses no interdependency, \*! - only addresses security to safety assessment]

Risk is an uncertain event [61], and such uncertainty is better represented and handled with probabilities [36, 35]. A push towards quantitative risk assessment is rising to represent risk better. Quantitative risk analysis attaches specific numerical values to determine risks. One significant advantage of using quantitative risk assessment is that one can change the format of the results in monetary form or in the form one wants. The challenges involved in quantitative risk assessment are sharing security-relevant information, insufficient historical data, complexity in knowledge transfer, and the dynamic nature of risk [36].

We believe quantifying risk values is crucial for automating risk assessment as it provides standardized metrics for consistent evaluation and prioritization [15]. Numerical risk values enable data-driven decision-making, allowing automated systems to prioritize mitigation strategies and allocate resources efficiently. Quantified data serves as input for models, enhancing predictive capabilities and refining risk models over time.

## 2.3 Risk Treatment

Risk treatment is a vital component of risk management in complex systems, particularly in ICSs, where both safety and cybersecurity concerns coexist. The objective of risk treatment is to ensure system operation remains resilient, secure, and sustainable under adverse conditions. A key guiding framework for this process is the ALARP (As Low As Reasonably Practicable) principle, which aims to reduce risk to a level where further reduction is impractical due to cost, time, or technical constraints [3].

The four primary strategies of risk treatment are mitigation, avoidance, transfer, and acceptance [11]. Depending on the measures usually in place in the organization, one or more of the four routes are considered. The research mainly focuses on the mitigation route based on the risk identification and assessment results. The choice of a route depends on factors like the economic aspect, technical capability, available tools, nature of the attack, and third-party contracts of the organization [3].

The risk treatment phase depends on the assessment. In the assessment, safety hazards and security vulnerabilities are identified. The risk treatment should combine hazard-based and vulnerability-based mitigations [17, 24]. The method used for risk assessment and management should incorporate a safety and security approach.

In [98], even identification of the hazard is considered risk treatment. However, it is the first out of the six levels identified for risk treatment. Other levels include: *identification of worst-case scenario, plausible worst cases, best estimate, probability analysis, and display of risk uncertainties*.

Organizations employ risk mitigation strategies beyond hazard-based, vulnerability-based, and threat-based approaches. Resilience-based strategies build organizational adaptability and redundancy, while deterrence-based strategies discourage threats through visible deterrents. Redundancy in critical systems and processes mitigates risks by ensuring backup mechanisms. Insurance and risk transfer involve transferring certain dangers to third parties.

### 2.4 Risk Management

Risk assessment is an important part of identifying such risks beforehand through the knowledge of failures, threat vectors, and vulnerabilities. It is usually *Agent-based* or *Business-based* [3]. Risk management, however, recommends some mitigation techniques for attacks. Risk management provides important insight into the safety and security implementation of industries. For initial risk assessment, [55] suggests performing an analysis without security postures. This will help the organization gain in-depth knowledge of the required steps to follow for safety and security implementation. As mentioned in [60], risk management systems must follow certain characteristics. Risk management and methods involved should,

- (a) create value (human health and safety, product quality, financial value),
- (b) be an integral part of organizational processes,
- (c) be part of decision making (decision w.r.t. risk identification and treatment),
- (d) explicitly addresses uncertainty (unknown situation),
- (e) be systematic and structured,
- (f) be based on the best available information,
- (g) be tailored,
- (h) take into account human factors (internal and external risk actors),
- (i) be transparent and inclusive,

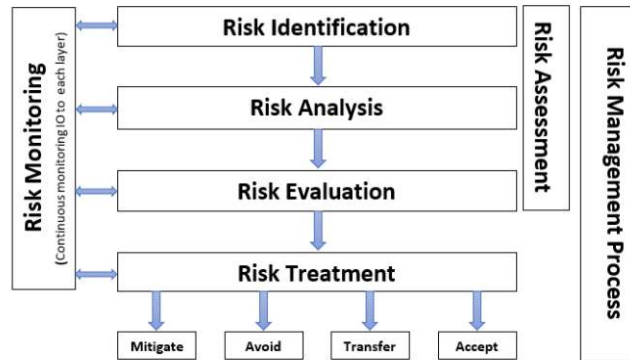


Figure 2.1: Risk management process

- (j) be dynamic, iterative, and responsive to change, and
- (k) be capable of continual improvement, update, and enhancement.

According to [14], the primary focus of risk management should be to

- identify the risks with respect to the provided safety and security to personnel, process, information, and assets,
- analyze the damage to the environment and image (reputation) of the organization in times of an accident or attack, and
- evaluate the stability and maintenance of the overall equipment efficiency of components.

Mitigation steps for safety and security are covered in standards and best practices ([60], [122], [55], [66]). In [4], a defense-in-depth approach for securing ICSs is mentioned. Figure 2.1 shows an essential risk management flowchart highlighting the steps. The As Low As Reasonably Practicable (ALARP) principle is a common risk-reduction principle based on risk-informed and cautionary/precautionary thinking [11].

## 2.5 Information Model for Risk Assessment-Single Source of Truth

The ontological approach has emerged as a leading and widely endorsed method for addressing safety and security risk assessment, as demonstrated by a substantial body of research [12, 87, 9, 92, 75, 127, 126, 101, 7, 139, 17]. Ontologies offer a formal, semantically rich framework that enables consistent risk representation, reasoning, and interoperability

across systems, making them particularly well-suited for complex, multi-domain industrial environments.

The other information models like AML and AAS have also gained popularity. These models are being actively explored for their potential in system modeling, data exchange, and digital integration, contributing to enhanced risk assessment capabilities. In industrial settings, the primary function of AML is to represent domain-specific information using the Computer Aided Engineering Exchange (CAEX) format, as detailed in [147]. It is widely used for modeling systems in automation, communication, behavior, and processes. AML also enables data exchange across various engineering tools, both within its own structure and with external formats such as ontologies. Beyond interoperability, it supports conceptual modeling, innovative engineering workflows, and integration of diverse digital systems.

Initial efforts to apply AML in risk assessment are seen in [23], which introduces a model-based approach using tailored metrics to evaluate project risks from distributed engineering data, demonstrated through tool support and a real-world case. However, challenges such as handling complex, evolving projects and reliance on accurate data persist. [39] explores using AML artifacts to identify security risks and predict attack outcomes, while [37] extends this to analyze cascading quality impacts. Foundational research [16, 47, 33, 34, 23] underpins our use of AML as a data source for a BBN-based integrated risk assessment, though similar limitations—particularly regarding model complexity and data reliability—remain relevant.

The AAS serves as the digital twin of a physical asset and is developed under the Platform Industrie 4.0 initiative<sup>3</sup>. It structures asset-related information into modular sub-models encompassing technical specifications, operational data, maintenance records, and communication interfaces [13]. The AAS architecture includes three components: a passive data layer, an application programming interface (API), and an active component. Information storage and access are primarily handled by the passive layer via the API [14]. Several studies have demonstrated the use of AAS in digitizing individual assets and mini-factories [140, 116, 62, 118]. Its potential for risk assessment is discussed in [14], while the integration requirements for addressing both safety and security concerns through AAS are outlined in [52].

Our research indicates that both AML and AAS are employed for asset representation and risk assessment. AML excels in offering a standardized, detail-rich format for modeling and exchanging engineering data, particularly effective during the design and configuration stages. In contrast, AAS provides a more dynamic, lifecycle-oriented digital representation that includes real-time operational insights and historical maintenance data. The information models used for the single source of truth have been undertaken in the past. However, their use for an integrated safety and security risk assessment is demonstrated in this research. The actual design concept and implementation strategies are mentioned in the following chapters.

---

<sup>3</sup><https://industrialdigitaltwin.org/en/>

## 2.6 Human Errors in Risk Management

Human factors significantly contribute to accidents in ICSs, with errors such as action, checking, retrieval, transmission, diagnostic, and decision errors being prevalent [2, 48]. Organizations should identify root causes and mitigate errors through updated procedures, training, or system redesigns [124]. However, these error classifications lack a structured, model-based representation.

Human Reliability Analysis (HRA) research explores the impact of human error on safety, quality, productivity, and loss, as well as factors influencing errors [97, 142]. Performance-shaping factors, derived from methods like THERP, HEART, ATHENA, and SPAR-H, offer qualitative insights but lack quantitative rigor and practical guidelines [81].

BBNs, as quantitative probabilistic tools, address uncertainty and are applied in ICS for fault detection, accident analysis, and human error modeling [8, 123, 95, 88]. These models explore organizational factors, dependencies, situational awareness, and HRA extensions but lack integration of security-related incidents for comprehensive interconnected system analysis [134, 68].

BBNs estimate human error probabilities in domains like nuclear power, enhancing cognitive reliability [5]. Machine learning-based approaches also identify factors influencing errors, emphasizing cognitive, organizational, and technological aspects [91]. However, most HRA relies heavily on expert judgment due to limited data availability [97].

To address data scarcity, the Multi-Attribute Technological Accidents Dataset (MATA-D) has been used to estimate human error causes and effects, with BBN representing factor relationships [90]. Proposals include leveraging AI tools with similar datasets [64]. Simulated probabilistic data distributions, such as those used in SPAR-H, provide another avenue for estimating performance-shaping factors [41].

According to ISO 31000 [60], risk management is incomplete without incorporating the human factor, as it is a critical element influencing the likelihood and severity of risks in any system. Human factors encompass individual behaviors, decision-making processes, communication, and interactions with technology and organizational structures. The standard emphasizes that failing to consider human contributions, whether as potential sources of risk or as pivotal in risk mitigation, can lead to inadequate risk assessments and control measures. Integrating human factors into risk management processes enhances the ability to identify vulnerabilities, anticipate potential failures, and implement effective strategies for prevention and its recovery.

## 2.7 Automation process and levels

An automation system operates with minimal or no human intervention, executing repetitive tasks through rule-based responses in a structured environment [44]. Risk assessment relies heavily on human input, limiting its automation and posing significant

challenges. It also supports real-time analysis, enabling dynamic risk updates and responses to evolving conditions [63].

In [117], first principles and levels of automation are mentioned. The author provides ten levels of automation from complete manual to complete automation by the computer. Further analysis for automation levels for driving systems is provided by SAE <sup>4</sup>. They provide levels ranging from no automation in driving (Level 0) to full automation in driving (Level 5).

According to [15], the automation of risk assessment and management provides the organization with the following: *Automated prioritization for Risks, Ability to Incorporate Change, Addressing Dependency between Risks, Automated Risk Remediation calculation, and Automated Risk Report*. The automation levels considered for this research are referred to from the [44]. The authors propose six levels of autonomy for various plant processes. The evaluation of the model is based on mapping the levels to the actual performance.

As per our current knowledge, there has not been an integrated safety and security assessment methodology that is both semi-automated and capable of addressing the interdependencies among different elements of a system. Existing approaches tend to treat safety and security separately, and often lack mechanisms to model the dynamic interactions between various components, including technical systems, human factors, and organizational influences.

To date, no methodology has effectively bridged these aspects in a unified and semi-automated framework. However, the use of BBNs has been demonstrated in various publications as a practical and robust tool for risk assessment. BBNs are particularly advantageous due to their ability to handle uncertainty, represent causal relationships, and update beliefs as new evidence becomes available. Moreover, BBNs can incorporate expert judgment and empirical data, and are capable of integrating human error analysis, making them well-suited for complex systems where human interaction plays a critical role. Their graphical nature enhances transparency and facilitates communication among stakeholders, while their probabilistic foundation enables more informed and nuanced decision-making.

---

<sup>4</sup>[https://www.sae.org/standards/content/j3016\\_202104/](https://www.sae.org/standards/content/j3016_202104/)

# Safety, security, and their relationship

This chapter mentions the core definitions of safety, security, and their elements. Critical differences and similarities are highlighted to bring them together. It aims to understand the current status of academia and industry regarding safety and security research and awareness, respectively.

## 3.1 Safety

Safety is a condition that gives one freedom from hazards, risks, and accidents that may cause injury, damage, or even death to people, processes, and the environment [2]. Safety in industries is implemented to protect the workers from the danger of commercial accidents. Safety is of primary concern for industrial facilities, from protecting employees from accidental injuries to safeguarding equipment from severe damage, which may result in costly downtime and repairs. Industrial safety is the ability to determine the risks inherent to operations and safety functions [102]. It is a commitment to identify dangers in production operations, assess them in terms of quality and quantity, and manage them. Employees are obligated to the assigned tasks and will minimize any negative aspects of the situation affecting employees' health and safety [129]. A fault in an asset leads to an error in its function, resulting in failure. A prominent example of such a failure chain is a worn-out motor (fault) that does not produce enough revolutions per minute (error) and causes the function to fail (failure). Safety incidents such as hardware or software faults, human errors, or natural calamities lead to component failures [15]. Fig. 3.1 provides a conceptual view of safety in the industry.

Safety has long been a significant concern for organizations, especially with the advent of hazardous technologies and activities. Safety is a core concept in policy, regulation,

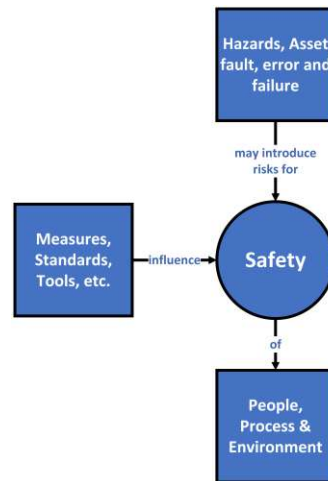


Figure 3.1: Safety in industry

and management in the energy, chemical, transportation, water, and health sectors [131]. Consequently, there are well-established industry strategies, collaborations, and practices for preventing incidents and accidents. Research demonstrated how hazards relate to changing organizational characteristics, and the argument that significant accidents are inevitable in specific high-hazard systems became influential and spurred interest in the limits to safety and possibilities of organizational competence [102]. According to data published by [2], humans are the weakest link in an ICS environment.

## 3.2 Security

Safety was given much attention and focus due to highly critical systems with significant impacts on their environments for a long time. However, only accidental component failures or software errors were traditionally addressed. Today, system safety can also be compromised by security breaches and electronic attacks [78].

Originally, ICS implementations were susceptible primarily to local threats because many of their components were in physically secured areas, and the elements were not connected to IT networks or systems [131]. However, integrating ICS components with IT networks provides very little isolation for ICS from the outside world than traditional systems. It creates a greater need to secure these systems from remote, external threats [40].

Also, the increasing use of wireless networking places ICS implementations at greater risk from adversaries in relatively close physical proximity, but who do not have direct physical access to the equipment. Threats to the ICS can come from numerous sources. It includes, but is not limited to, hostile governments, terrorist groups, natural disasters, and malicious or accidental actions by insiders [51].

Security attacks deal with vulnerabilities of assets exploited by threats. These vulnerabil-

ities, once exploited, increase the risk of unsafe operations. Over the years, the security attacks on ICSs have increased, as evidenced by prominent examples such as Stuxnet, Black Energy, WannaCry, and DarkSide [24].

Safety is generally related to component failures, decreased efficiency, failure causes, probability, severity, and unintentional hazards. In contrast, security is related to the development of intentional attacks, asset vulnerability, and criticality of the facility to hamper the performance or gather information through the physical or cyber-realm, enabled internally or externally [15]. Fig. 3.2 provides a conceptual view of the security parameter implemented concerning safety in the industry. It indicates that security is in place to ensure the safety of people, processes, and the environment.

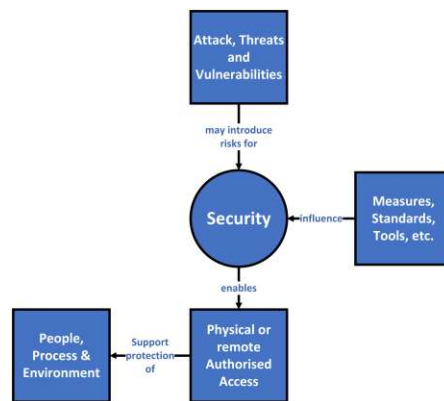


Figure 3.2: Security in industry

The implementation of the security measures is with the fulfillment of the following recommendations [122, 121]:

- Restricting logical access to the ICS network and network activity.
- Restricting physical access to the ICS network and devices.
- Protecting individual ICS components from exploitation.
- Maintaining functionality during adverse conditions.
- Restoring the system after an incident.

### 3.3 Similarities and differences between safety and security

From the definitions, one can argue that there are some similarities between safety and security. The first and foremost similarity is that safety and security aim to avoid risks and ensure smooth functioning of the ICS. In [78], commonalities concerning dealing with risks, resulting in constraints, involving protective measures, and creating requirements are mentioned. Some techniques that apply to one field can also be used in another.

The commonalities are also evident when we consider the input requirements. In [15], shared data sources such as system knowledge or asset information used for safety and security risk assessment are mentioned. However, the difference comes about from information usage. In [106], similarity in design and operation principles is noted for detailed monitoring, defense in depth, in-depth knowledge of the system, complexity, and maintaining detailed inventories.

The difference between safety and security, according to our understanding, can be visualised by the origin of risk: safety considers hazards (i.e., system failures). In contrast, security considers threats and focuses on how potential attacks may impact the system's assets and operations due to vulnerability.

The nature of consequences is also different: safety is related to risks that could potentially impact the system environment, while security is related to risks that can have consequences on the system itself or its environment [106]. Safety and security were isolated for a long time, as evident by differences in the tools, standards, and risk management in the two domains [14]. Fig. 3.3 provides a conceptual view of safety and security working together for the common goal.

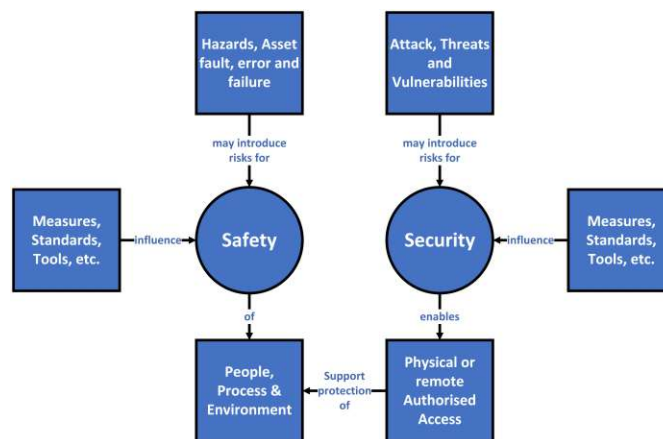


Figure 3.3: safety and security in industry

## 3.4 Relationship

- *Conditional Dependency*: Implementing safety measures may put some conditions on security or vice versa. For example, a safety measure for an incident (fire: open doors to evacuate) weakens security posture (reduced or no authentication or authorization). It may lead to a malicious attack (e.g., an attacker may physically enter the building with no authentication or authorization), increasing the risk.
- *Mutual Reinforcement*: Application of safety measures enhances security or vice-versa, enabling resource optimization and cost reduction. For instance, an Intrusion

Detection System (IDS) implemented as a security measure can be used to analyze the safety impact based on the target identified.

- *Antagonism*: When considered, safety and security requirements or measures lead to opposing situations. For instance, in an automatic door control, a safety measure would be to open the door in case of a safety alarm, while a security measure would be to close it in case of a security attack. Here, the function of the door behaves oppositely. As another example, consider the target response time. In the case of safety, the response should be immediate, while security measures may demand first authorization and then take care of safety measures, thus increasing the response time.
- *Independent*: No interaction at all. For instance, safety and security should be considered independently under normal working conditions. Fig. 3.3 provides a view of normal operation.

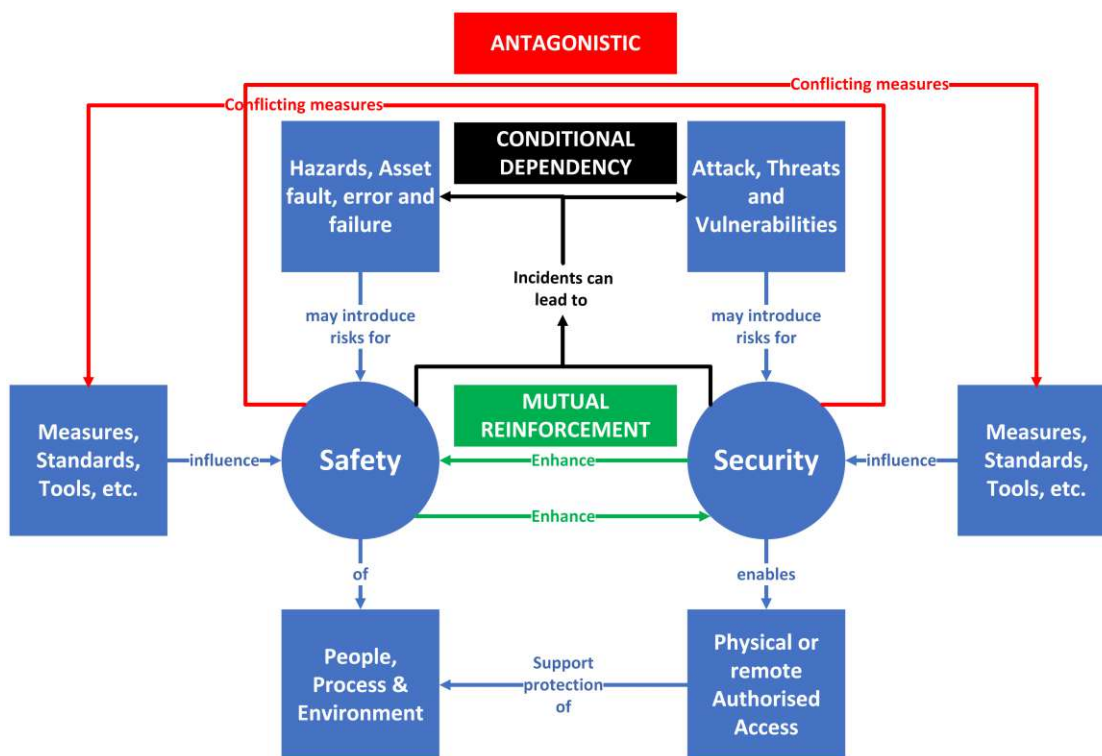


Figure 3.4: Safety and security relationship: Antagonistic, Conditional Dependency, Mutual Reinforcement, and Independent.

## 3.5 Benefits of integrating safety and security

Besides improving the safety and security of systems, integrating safety and security can lead to several benefits. In [96], some benefits of integrating safety and security are mentioned:

- **Early-On Integration of Safety and Security:** Safety and security assessments can be carried out while the requirements of system features are established. Safety assessments provide concrete hazards that should be treated by security assessments, thus helping security engineers to set priorities. For example, a safety hazard shall be prioritized over other security attacks that do not cause catastrophic events.
- **Verification and Validation:** While safety has many well-established methods for verification, security verification relies mainly on penetration testing techniques, which are system-dependent and, therefore, resource-intensive. The integration of safety and security can facilitate security verification. Much of the knowledge can be retrieved from safety assessments, thus saving resources. For example, FTAs describe the events leading to some hazardous event, while FMEAs describe single points of failure. Security engineers can use this information to plan penetration tests, e.g., exploit single points of failure described in FMEAs, thus leading to increased synergies and fewer development efforts.
- **Safety and Security Mechanisms Trade-Off Analysis:** By integrating safety and security analysis, it is possible to analyze trade-offs between control and countermeasures proposed to support safety and security arguments. On one hand, safety and security measures may help each other, making one of them superfluous. For example, there is no need to use CRC (Cyclic Redundancy Check) mechanisms for safety if messages are being signed with MAC (Message Authentication Codes), as the latter already achieves the goal of checking for message corruption. On the other hand, safety and security mechanisms may conflict. For example, emergency doors increase safety by allowing one to exit a building in case of fire, but they may decrease security by allowing unauthorized persons to enter the building. Such trade-off analysis can help solve conflicts and identify and remove redundancies, reducing product and development costs.

# Methodology

This research follows the design science research process (DSRP) [100], adapted to address the objectives of developing an automated and integrated safety and security risk management system. Fig. 4.1 illustrates the modified DSRP framework aligned with the research goals.

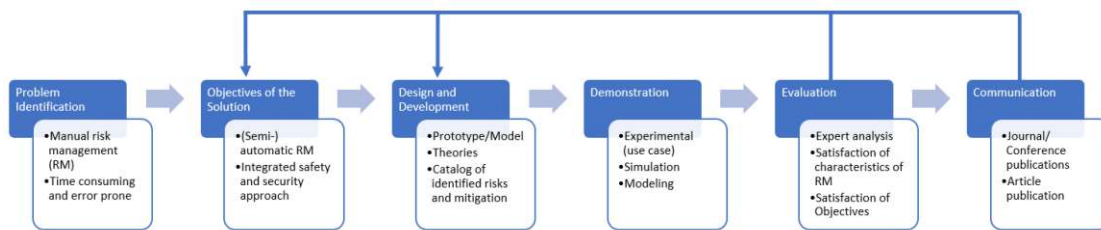


Figure 4.1: DSRP modified for current research [100]

## 4.1 Problem identification and motivation

This phase establishes the significance of the research, highlighting the challenges of current manual risk assessment methods, which are time-consuming and prone to errors. Also, an integrated safety security approach to the risk assessment in ICS is needed. The Sections 1.1 and 1.2 in chapter 1 highlight the problem statement and the motivation for the research. The research aims to address the inefficiencies by creating a system that integrates safety and security considerations. The solution aligns with advancements in I4.0 and the IIoT to meet future technological demands.

### 4.2 Objectives of the solution

The objective of the research is to satisfy the previous phase. The proposed solution will help to develop a (semi-)automatic risk management model/prototype, highlighting the following:

1. Identification of safety and security needs of the organization.
2. Identify hazards, threats, and vulnerabilities of the SUC considering safety and security interdependencies.
3. A model/prototype to analyze and evaluate risks.
4. Suggest mitigation and support decision-making.

The Section 1.3 in Chapter 1 provides a comprehensive overview of the research objectives, offering clarity on the goals and scope of the study. It also presents the research questions instrumental in guiding the investigation. The primary focus is on systematically addressing these questions to achieve the intended objectives while ensuring alignment with the broader purpose of the research. This foundation establishes a clear trajectory for the subsequent chapters and methodologies, ensuring coherence throughout the study.

### 4.3 Design and development

To achieve the objectives, this phase focuses on creating models, simulations, and systematic procedures:

- **Asset-Based Risk Assessment Strategy:** An asset-centric approach will identify risks using various data sources (System Architecture, Asset Inventory, Documents for Hazard analysis and Vulnerability data, or Historian) and input from engineers or operators in the form of Interviews. This forms the basis for developing and maintaining a risk database for future updates.
- **Information Collection and Retrieval:** An information model stores and uses the information gathered from various data sources. This model serves as a structured framework for organizing and retrieving data efficiently. It integrates diverse types of information, such as historical data, real-time inputs, and expert knowledge, into a unified system. The model ensures consistency, facilitates data sharing across modules, and enables scalable information management. The information models development using AML, AAS, and Ontology are demonstrated (Chapter 6).
- **Quantitative Risk Assessment:** Safety and security risks will be evaluated using a quantitative approach, factoring in economic and physical consequences.

- **Integrated Risk Assessment:** Interdependencies between safety and security risks will be analyzed using a probabilistic framework. Out of different risk assessment methods compared, BBN is ideal for an integrated safety and security approach. As a probabilistic and graphical tool, it can be used for modeling interdependencies between assets, combining expert knowledge and empirical data, simulating "what-if" scenarios, providing diagnostic clarity for faults and their sources, offering dynamic adaptability for real-time updates, and providing risk propagation for cause and effect analysis. BBN's superiority over neural networks for specific tasks, such as fault diagnosis and source tracing, further supports its use. The method's practical challenges, merits, and demerits are evaluated.
- **Automated Risk Assessment:** Based on the information collected, stored, and retrieved, the developed prototype will perform the risk assessment. Any data or system update change shown in the information model will automatically be reflected in the risk assessment.

## 4.4 Demonstration

The efficiency of the previous phase is measured in this phase. Demonstration of the prototype of the semi-automated risk assessment and management requires a use case. The information related to the system is acquired using the use case. This information is then stored in the information model and used to develop the graphical network. The graphical network then provides the risk propagation for the whole system.

### 4.4.1 Use Case

Modular Production System (MPS 403-1) is an ICS representation of a miniaturized modular production line and offers deep insight into the intelligent networking of machines in the production environment. The system consists of three stations: distributing, joining, and sorting. For demonstration purposes, the focus of the modeling process is on the sorting station, which is enhanced with the integration of a collaborative robot (cobot). The cobot is operated through a dedicated interface. The primary function of the Sorting Station is to classify workpieces based on specific attributes such as color and type.

Fig. 4.2 illustrates the cyclic sorting process. The Sorting Station operates through a series of steps, comprising modules such as a conveyor, detection unit, sorting mechanism, and pick-and-place system. The process begins when the operator places a workpiece on the conveyor belt. An optical sensor detects the workpiece's presence and activates the conveyor belt. The belt transports the workpiece to the detection module, a specialized unit that analyzes the workpiece's attributes.

The detection module evaluates features like presence, color, and type. Following detection, the workpiece moves to the sorting module, where solenoid deflectors guide its trajectory based on the detection results. For instance, red or black workpieces with

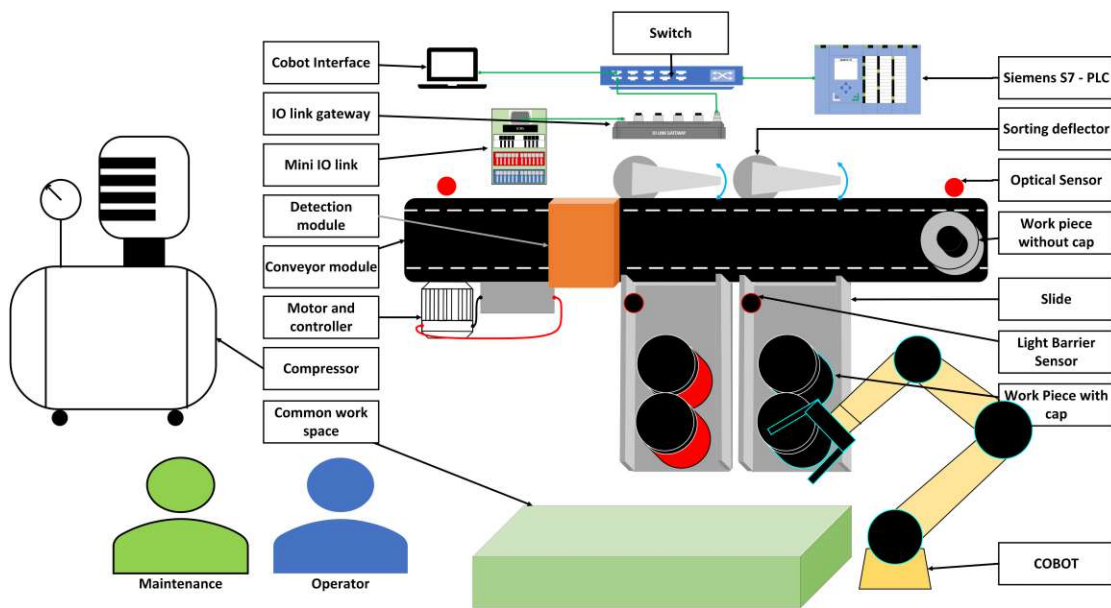


Figure 4.2: Connection and components of the Sorting Station

caps are sorted into designated areas, while silver workpieces without caps continue to the end of the conveyor belt.

The cobot executes its programmed task (programmed by cobot controller (CBC)) of picking workpieces from slides or the conveyor and placing them on the shared workspace according to the detection outcomes. The cycle concludes when the operator collects the workpiece from the shared workspace and places it back on the conveyor for the next iteration. Maintenance personnel handle routine upkeep, such as repairing damaged parts and ensuring system functionality. The process is automated using a PLC, along with sensors, actuators, and communication devices, ensuring efficient and consistent operations.

#### 4.4.2 Development of Prototype

The demonstration of the prototype is necessary for evaluation and making further changes. To develop the prototype, we use a Python program. Python provides us with the required libraries for accessing the information stored in information models, retrieving the data, developing a network graph, and providing the analysis using BBN.

Developments of theoretical methods for automating potential measures (technical and organizational) for risk reduction were researched and applied or suggested. The methods practical for implementation are documented and modeled, and a step-by-step implementation of developed concepts in the lab environment is done to make a note of their effectiveness and functionality.

## 4.5 Evaluation

The developed prototype model is evaluated using analysis by both safety and security experts. The requirements provided by the standard ISO 31000 for the risk management method are also compared for evaluation. The use case demonstration offers the basis for the review.

## 4.6 Communication

This step is the final phase, during which the results will be documented. The gained results will be published in relevant journals/conferences. The results of the research are available in the following publications.

1. P. Bhosale, W. Kastner and T. Sauter, "A Centralised or Distributed Risk Assessment using Asset Administration Shell," 2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA ), Vasteras, Sweden, 2021, pp. 1-4, doi: 10.1109/ETFA45728.2021.9613152.
2. P. Bhosale, W. Kastner and T. Sauter, "Automating Safety and Security Risk Assessment in Industrial Control Systems: Challenges and Constraints," 2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA), Stuttgart, Germany, 2022, pp. 1-4, doi: 10.1109/ETFA52439.2022.9921517.
3. P. Bhosale, W. Kastner and T. Sauter, "AutomationML use for Safety and Security Risk Assessment in Industrial Control Systems," 2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA), Sinaia, Romania, 2023, pp. 1-4, doi: 10.1109/ETFA54631.2023.10275476.
4. P. Bhosale, W. Kastner and T. Sauter, "Integrated Safety-Security Risk Assessment for Production Systems: A Use Case Using Bayesian Belief Networks," 2023 IEEE 21st International Conference on Industrial Informatics (INDIN), Lemgo, Germany, 2023, pp. 1-6, doi: 10.1109/INDIN51400.2023.10217926.
5. P. Bhosale, W. Kastner and T. Sauter, "Integrated Safety-Security Risk Assessment for Industrial Control System: An Ontology-based Approach," 2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA), Sinaia, Romania, 2023, pp. 1-8, doi: 10.1109/ETFA54631.2023.10275530.
6. Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. 2024. Modeling Human Error Factors with Security Incidents in Industrial Control Systems: A Bayesian Belief Network Approach. In Proceedings of the 19th International Conference on Availability, Reliability and Security (ARES '24). Association for Computing Machinery, New York, NY, USA, Article 53, 1–9. <https://doi.org/10.1145/3664476.3670875>

7. B. Brenner, S. Hollerer, P. Bhosale, "Better Safe Than Sorry: Risk Management Based on a Safety-Augmented Network Intrusion Detection System," in IEEE Open Journal of the Industrial Electronics Society, vol. 4, pp. 287-303, 2023, doi: 10.1109/OJIES.2023.3297057.
8. P. Bhosale, W. Kastner and T. Sauter, "Mapping ICS Vulnerabilities: Prioritization and Risk Propagation Analysis with MITRE ATT&CK Framework and Bayesian Belief Networks," 2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA), Padova, Italy, 2024, pp. 1-8, doi: 10.1109/ETFA61755.2024.10710893.
9. P. Bhosale, W. Kastner and T. Sauter, "Comparative Analysis of AAS and AML as a Data Source for Integrated Risk Assessment in ICS," 2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA), Padova, Italy, 2024, pp. 1-4, doi: 10.1109/ETFA61755.2024.10711039.
10. P. Bhosale, W. Kastner and T. Sauter, "AutomationML Meets Bayesian Networks: A Comprehensive Safety-Security Risk Assessment in Industrial Control Systems," in IEEE Open Journal of the Industrial Electronics Society, vol. 5, pp. 823-835, 2024, doi: 10.1109/OJIES.2024.3439388.

# Information Collection Phase

In today's digital era, data is a critical asset for organizations. Proper data collection is essential for improving efficiency, making informed decisions, and developing data-driven algorithms. It involves gathering and organizing relevant information to support business decisions and is crucial for feature extraction and implementing semi-automatic risk identification and management schemes. Organizations use various data collection methodologies to gather the necessary inputs. Despite these methods, challenges in understanding and extracting information persist, often requiring human-user interaction to address them.

The data available within ICS includes both engineering and operational (runtime) information. Engineering data provides insights into the system's construction and design, while operational data reveals how the system functions during runtime. This chapter identifies the necessary data from various sources and its use in risk assessment. The data needed from multiple sources and some techniques to get that information are also mentioned.

## 5.1 Data Input Sources

### 5.1.1 Interview and Questionnaire

Interviews with system users and management give risk assessors critical insights into system operations. On-site visits offer a firsthand understanding of the physical, environmental, and operational safety and security aspects. This step is a key component of the overall data collection process. While primarily qualitative, the information gathered can include both qualitative and quantitative data.

The questionnaire aims to collect information pertinent to business requirements and the application of relevant safety and security standards (if implemented). It is designed for

technical and non-technical management personnel designing and operating the system. This method of data collection can be conducted online or offline. Examples of typical questions include:

- What are your assets?
- What hazards are identified for operations?
- What threats to those assets are known?
- Which security attributes (confidentiality, integrity, availability, etc.) of them are vital?
- What are the restrictions and demands of the stakeholder (e.g., does he have requirements on risks from a legal or compliance perspective)?
- How often do you access and update your risk management program? When was the last risk management performed?
- What strategies are implemented in your organization to address (identify, assess, and manage) risks?
- If disaster X occurs, which functions are likely more risky? What are the likely and possible scenarios and expected consequences?
- Where specifically is the information processed and stored?
- What is the percentage of human error that leads to risks?
- Based on a 90% confidence interval, what is the probability of X (something) happening? (These questions are for expert opinion when collected data is insufficient.)

### 5.1.2 Document Review

Documents as input are received from the management after getting the required approval. Experts can review documents to identify the required information. This can be a regular input whenever there is a change or update in the system. For instance, the documents reviewed are:

- Implemented policies (legislative, directives) - for asset utilization.
- System documentation (manuals, design, and architecture) - for the system user guide.
- Security related (previous risk assessment, recommendations) - for threats and vulnerabilities.
- Safety-related (hazard analysis) - for hazard identification.

Documents offer valuable insights into the system and business requirements and facilitate safety and security analyses. Additionally, documents detailing asset information, such as asset lists, the number of assets in use, and electronic data sheets, are essential. While document review is unsuitable for automating data collection, the input is typically required only once or when changes occur. Human users can manually input the necessary data as needed.

### 5.1.3 Automatic data gathering tools

Automatic data-gathering tools enable efficient and accurate information collection, supporting real-time and historical data analysis. These tools provide insights into services running on the host or network, and their outputs are invaluable for identifying vulnerabilities and performance issues. Tools such as IDS, network scanning utilities, and Security Information and Event Management (SIEM) systems are beneficial for monitoring security events, analyzing network traffic, and detecting anomalies.

In the ICS environment, data is stored in databases designed for high performance and reliability, such as SQL-based, NoSQL-based, or time-series databases. These databases, called Data Historians, capture information on asset performance, process statuses, efficiency metrics, and failure probabilities. Historical data from these systems aids in analyzing past incidents, asset impact history, and the likelihood of future issues.

The retrieval of stored data is streamlined through client-side functions, ensuring seamless access for analysis and decision-making. Data transfer is facilitated using modern technologies, including Application Programming Interface (API), integrated cloud-based services, and specialized data drivers. Furthermore, IIoT protocols like Open Platform Communication unified Architecture (OPC UA) and Message Queuing Telemetry Transport (MQTT) play a critical role in securely transferring or publishing data. These protocols allow clients or subscribers to access only the data permitted by the server or broker, maintaining both accessibility and data privacy.

### 5.1.4 Online databases

Risk managers use online databases to get information about vulnerability (existing, exploited), asset-related, and proprietary information (MTBF testing by the company). The databases are mostly free and have minimal restrictions for usage. It has proven to be reliable and has been used by researchers and engineers alike. Users can use automatic data-gathering tools to obtain updated information from these websites. Online databases include:

- *National Vulnerability Database (NVD)*<sup>1</sup>: The NVD includes databases of security checklist references, security-related software flaws, misconfigurations, product names, and impact metrics.

<sup>1</sup><https://nvd.nist.gov/vuln/search>

- *ICS CERT Cybersecurity & Infrastructure Security Agency (CISA)*<sup>2</sup>: It is a database maintained by the US government that provides reports for cyber incidents and exploited vulnerabilities.
- *MITRE*<sup>3</sup>: A US non-profit that supports government agencies and hosts CVE, an international cybersecurity collaboration.
- *Microsoft security bulletins*<sup>4</sup>: A library of security documents released by the Microsoft Security Response Center.
- *BugTraq SecurityFocus*<sup>5</sup>: It is a full disclosure mailing list for full disclosure of security vulnerabilities for IT systems.
- *Vendor Provided Vulnerabilities*: Data provided by OT vendors concerning disclosed vulnerabilities in their products.

The generic publishers mentioned above (NIST, ICS-CERT, and MITRE) make multiple formats available for downloading Common Vulnerabilities and Exposures (CVE) information, including CSV, XML, JSON, text, and HTML. Data is updated frequently, often hourly, and should be synced regularly to an asset inventory. Downloaded CVEs can be matched with inventoried assets using Common Platform Enumeration (CPE) product identifiers to check for vulnerabilities. Process owners are responsible for updating asset inventories with vulnerability disclosure information per OT device (the one that is connected).

The inputs chosen for risk management are sourced from various stakeholders, including human users (developers, operators, risk managers, and maintenance personnel), historians, and automated tools (e.g., IDS, SIEM). These inputs encompass system architecture, asset inventory, lists of utilized assets, previous risk assessments, misuse incidents, and historical data. Such information is critical in enabling humans to support and actively participate in risk management.

### 5.1.5 Vulnerability Scan tools

A vulnerability scan was conducted as part of a security evaluation of a Modular Production System (MPS), focusing primarily on the Sorting Station, one of the three identified subsystems. The goal was to assess potential cybersecurity risks associated with the PLC, the collaborative robot controller (CBC), and other connected components within this module.

Vulnerability, as per IEC 62443, is defined as a flaw or weakness in the system's design, implementation, or operation. A component can be vulnerable by itself or make the

---

<sup>2</sup><https://www.cisa.gov/>

<sup>3</sup><https://attack.mitre.org/matrices/ics/>

<sup>4</sup><https://learn.microsoft.com/en-us/security-updates/securitybulletins/securitybulletins>

<sup>5</sup><https://seclists.org/>

system vulnerable with some combination. The vulnerability of the system or component can be scanned by various tools available in the market. Tools such as Nmap, OpenVAS, or Nessus Essentials are freely available.

The system was first accessed physically via an Ethernet connection to the local network. An ARP (Address Resolution Protocol) scan was performed to identify active devices and their respective IP addresses. It revealed a list of connected components within the network segment. Subsequently, Nmap<sup>6</sup> was used to perform a detailed scan on the identified IP addresses, which helped to enumerate open ports and detect the services running on each device.

Following this reconnaissance phase, a vulnerability assessment was conducted using Tenable Nessus Essentials<sup>7</sup>, a free vulnerability scanning tool. Nessus Essentials enabled the identification of CVEs associated with the discovered services and firmware versions. NIST<sup>8</sup> provides a database of known vulnerabilities. The results included details such as the CVE identifier, Common Vulnerability Scoring System (CVSS)<sup>9</sup>, and potential mitigations. The CVSS is a framework used to quantify the severity of software vulnerabilities. Further information on CVSS is provided in Section 7.4.2. In this assessment, we use CVSS v3.1, which scores vulnerabilities on a scale from 0 to 10, with higher scores indicating more severe issues. Table 5.1 lists the vulnerabilities and their CVSS string and score for the assets.

ID	Name	Asset	CVSS String	Score
V1	Denial of Service (DOS)	CBC	CVSS:3.1/AV:A/AC:L/PR:N/UI:N/S:U/C:N/I:N/A:H/E:P/RL:T/RC:R	6.5
V2	BruteForce	CBC	CVSS:3.1/AV:A/AC:H/PR:L/UI:N/S:C/C:L/I:L/A:L/E:P/RL:T/RC:R	5.5
V3	Privilege Escalation	CBC	CVSS:3.1/AV:A/AC:H/PR:L/UI:R/S:C/C:L/I:L/A:L/E:P/RL:T/RC:R	5.1
V4	DOS	PLC	CVSS:3.1/AV:N/AC:L/PR:N/UI:N/S:U/C:N/I:N/A:H/E:U/RL:T/RC:U	7.5
V5	Code Injection	PLC	CVSS:3.1/AV:P/AC:L/PR:N/UI:N/S:U/C:H/I:H/A:H/RL:U	6.8

Table 5.1: Identified vulnerabilities from scan

<sup>6</sup><https://nmap.org/>

<sup>7</sup><https://www.tenable.com/downloads>

<sup>8</sup><https://nvd.nist.gov/vuln/search>

<sup>9</sup><https://www.first.org/cvss/>

### 5.2 Automation possibility of sources

The collected data can offer details about hardware, software, system interfaces, responsible users, functional requirements, and system limitations. Asset inventories provide information on the availability of assets. Previous risk assessments help identify existing safety and security measures, vulnerabilities, threats, and risks. Misuse incident records highlight potential entry points for threats, while historical data sheds light on normal system behavior and the impact of past safety or security incidents.

Some inputs must supply data continuously, while others may only be required periodically or in response to changes. Table 5.2 organizes this relevant information as follows:

- **Inputs:** Represents the data sources essential for individual risk assessment processes.
- **Data Collected:** Summarizes the type of data gathered from each input.
- **Frequency:** Indicates how often the data is required—continuously, periodically, or only during specific updates. It helps identify which inputs require automation for efficient assessment.
- **Digital Format:** Confirms that all inputs are stored digitally, facilitating easy access and sharing. "Digitization" here refers to data stored digitally, such as scanned documents or records maintained in spreadsheets.

The digitized data is often stored in databases, either locally or on the cloud, and is accessible following appropriate legal protocols. This ensures that the information is available to relevant stakeholders across the organization as needed for practical risk assessment and management.

### 5.3 Limitation or challenges

Although the benefits of choosing these sources are enormous and most risk assessment requirements are fulfilled, we must address some limitations and challenges with respect to the data collection methodology.

- **Necessary approval:** Every data collection source will require necessary approvals from top officials in charge of the operations.
- **Time:** In risk management, data collected from historians or SCADA can be too late to act upon a possible intrusion or security breach. Hence, getting alarms or notifications from the IDS or SIEM system of the detected possible positives is advised.
- **Bandwidth:** Although collecting data from historians or SCADA does not hamper their performance, the user's bandwidth limits this.

<b>Inputs</b>	<b>Data Collected</b>	<b>Frequency</b>	<b>Digital Format</b>	<b>Automation</b>
<b>Interview</b>	1) & 2) Expert opinion, experience, knowledge, etc.	One time	Available with expert support	Absent
<b>System Architecture</b>	1) Description of site architecture, safety barriers, connections, etc. 2) Network segments, security zones, connections, access points, etc.	One time or with updates in the system	Available with designer support	Absent
<b>Asset Inventory</b>	1) List of assets, component availability, etc. 2) List of assets to be protected, quantity, and vulnerabilities	One time or with updates in assets	Available with maintenance support	Present
<b>Process Definition</b>	1) Description of operations and processes, normal functions, parameters, etc. 2) Description of operations and processes, anomalies, normal operations, etc.	One time or with updates in processes	Available with developer support	Absent
<b>Documents and Policies</b>	1) Documents related to previous assessments, assets, safety procedures for system users, data sheets, standards, and regulations. 2) Documents related to previous assessments, assets, security postures, standards, and regulations	One time or with updates in the system	Available with management support	Absent
<b>Tools and Applications</b>	1) Historical data, past incidents, asset performance, component/process status, efficiency, failure probability. 2) Historical data, asset impact history due to security incidents, IDS, network scanning tools, and SIEM.	Continuous collection	Available with permissions	Present

\* [1: Safety, 2: Security]

Table 5.2: Identified Inputs for Safety and Security Risk Assessment [132, 15]

- Location: As data collection is possibly carried out remotely, the bandwidth limitation also depends on the user's location (remote or within the factory).
- Incomplete data collection: With bandwidth and location restrictions and a lack of approval for necessary sources, there can be a case of incomplete data collection or poor data quality. Such scenarios can lead to poor decision-making in the automated system, increased operational costs, lack of user visibility, etc.

## 5. INFORMATION COLLECTION PHASE

---

- Safety: The remote collection of data can impact the safety and intended use of the data. It needs to be secured.
- Security: Data safety is one of the primary concerns of any organization. A secure way to transfer data from one location to another is needed. Secondly, one more machine (laptop, computer, or data-storing device) is in use, making it a security concern for the organization.

The collected information is organised in the information models such as AML, AAS, and Ontology for further use in the automated risk assessment strategy. The information models serve as the single source of truth for the assessment process.

# Information Organization Phase

In this chapter, we identify different information models to organise different data. We implement the same use case in other information models. We also discuss how information models like AML, AAS, or ontology can provide a basis for data storage and be used for automating risk assessment. We identify and compare the models to determine the single source of truth input for risk assessment.

## 6.1 Automation Markup Language

AML is an Extensible Markup Language (XML)-based object-oriented data modeling language designed for storing and exchanging system engineering data. It is standardized under IEC 62714 and available as an open standard [33]. AML facilitates the exchange of engineering data across disciplines by describing assets as objects with varying aspects. For example, it can represent components like screws, grippers, robots, or entire production cells at different levels of detail. Objects are hierarchically structured, with smaller components linked to larger systems. AML integrates various standards through strongly typed links [34]:

Topology/Structure: Attributes and relationships of objects in their hierarchical plant structure, implemented using CAEX (IEC 62424).

- Geometry: Graphic attributes and 3D information, implemented with COLLADA by the Khronos Group.
- Kinematics: Connections and dependencies of objects for motion planning, implemented with COLLADA.
- Logic: Sequences, internal behavior, and I/O connections, implemented with PLCopen XML.

The development of AML is driven by the digitization initiatives of I4.0, aiming to streamline system documentation and improve the efficiency of engineering data exchange. It enables the detailed representation of system structure, behavior, and properties, addressing bottlenecks in the data flow to enhance engineering quality and efficiency [33].

AML has been widely adopted across industries for data storage and exchange. Examples of its use include creating comprehensive information models for automation components, exchanging data between engineering tools for material handling, automating project configurations between ECAD and PLC, modeling SCD diagrams (IEC PAS 63131), exchanging data for complex communication networks, and aligning engineering data with AAS specifications. These examples highlight its versatility in supporting modern industrial applications.

The core of AML is the top-level data format CAEX, which enables the structuring of the plant's engineering information. Based on XML, CAEX depicts physical or logical plant components in the form of data objects organized and described fundamentally by five modeling elements [33]. The same is mentioned in Fig. 6.1.

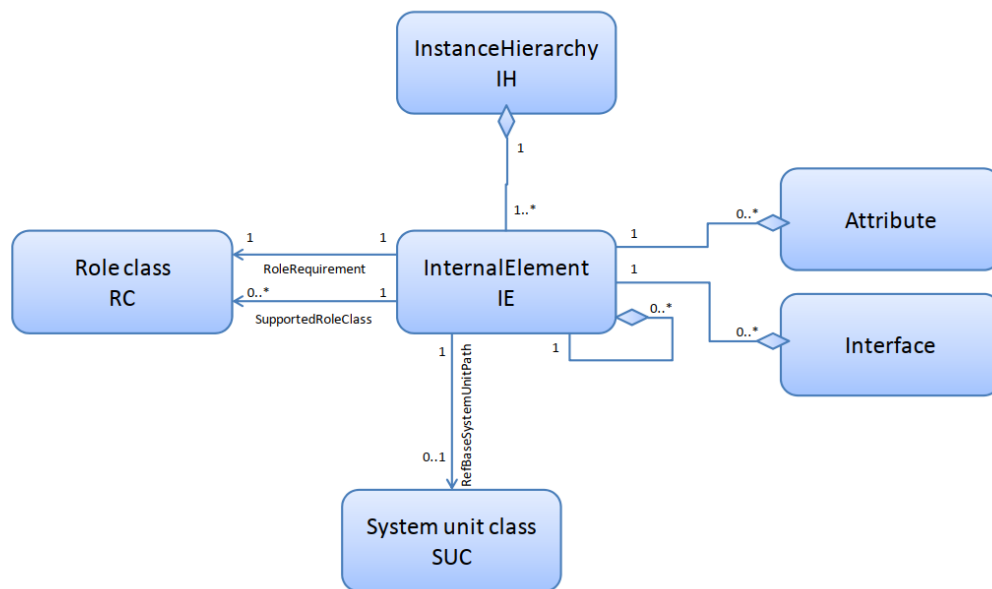


Figure 6.1: AML Simplified structure of Instance Hierarchy [33]

The fundamental elements of CAEX shown in Fig. 6.1 are defined as follows:

1. Instance Hierarchy (IH): IH describes the whole system and its specific hierarchy of components from top-level complete plant or functional component down to a single element like conveyor belt, wire, or PLC program (termed as Internal Elements (IEs)) with connection points (External Interfaces (EIs)) and related

connections (Internal Links (ILs)). It represents the engineering data modeled by CAEX following an object-oriented and hierarchical structure.

2. System Unit Class Library (SUCL): SUCL are sets of reusable modeling elements called System Unit Class, which define component types down to their respective technical realizations.
3. Role Class Library (RCL): RCLs are reusable role classes (RCs) for describing component requirements and semantic information, and referencing roles for IEs are used to identify their semantics.
4. Interface Class Librarys (ICLs): Reusable Interface Classes (ICs) for specifying connection points of RCs, SUCs, and interface types of IEs.
5. Attributes: Properties for characterizing each previously introduced modeling element.

### 6.1.1 AML model of Use case

The *IH* represents the actual components used in the experiment. It also represents the characteristics as *IEs or EIs or ILs* of certain components having different roles from *RCs* and classes from *SUCLs* and *ICLs*. The sorting station is divided into modules. The *IH* of the model is also divided into modules. Fig. 6.2 provides an instance on the *IH*, *SUCL*, *RCL*, *ICL*, and attributes. The model is developed in the AML editor <sup>1</sup>. The procedure is as follows:

1. Define the project structure based on the aim and use case
2. Develop SUCL: The whole project can be developed in *IH*, however SUCL help in instances that will be repeated.
3. Create instances of AML objects such as Systems, Devices, Components, Interfaces, etc.
4. Configure the properties and attributes (object names, object IDs, data types, descriptions, etc.) of each object in your AML model.
5. Define the interfaces representing the inputs, outputs, and behavior of the objects in the model.
6. Establish connections and relationships between the objects by linking interfaces between objects, representing data flows, dependencies, and interactions.
7. A validation and verification tool is provided in the editor to ensure the model follows the relevant standards and guidelines.

<sup>1</sup><https://www.automationml.org/download-archive/>

- The model can then be used for BBN development using the Python import command

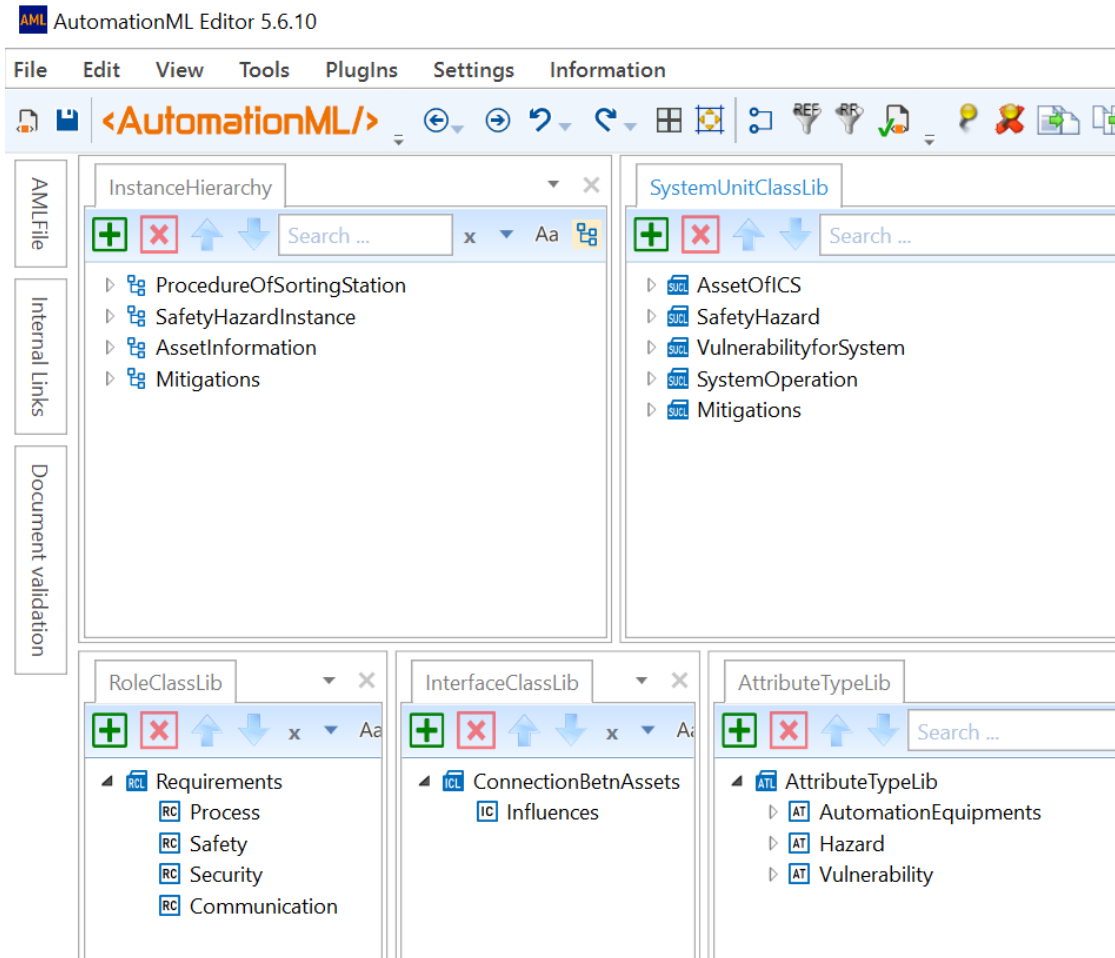


Figure 6.2: AML model for the use case: Sorting Station (Fig. 4.2)

### System Unit Class Library for the Sorting Station

The SUCL of the overall system is shown in the Fig. 6.3. It tries to capture the assets involved in the ICS operation. It helps instantiate the components of the targeted ICS system. Drag and drop the required classes in the IH. It contains information like the properties of the component it instantiates. Overall, SUCLs helps developers reduce the work of instantiating similar property assets used in the industry. The SUCL is enhanced with the ICL's instances.

In the model, the SUCL consists of categorizations such as *Hardware*, *Software*, and *User*. *Hardware* is further divided into *Process device* (contains automation devices used for

the process, such as sensors, actuators, controller, and workstation), *Machine* (contains valuable equipment for the process), and *Networking devices*. Each of these classes refers to the individual components to produce a greater granularity of information. *software* focuses on the applications, process logic, etc., while *User* is the human operator or maintenance personnel engaged in the system's operation.

### Instance Hierarchy of the Sorting Station

The AML model IH library shown in Fig. 6.4 of a sorting station is designed to automate the process of sorting workpieces based on their color, specifically red and black workpieces. The model consists of several modules working together to achieve this task effectively.

The IH library for the sorting station is divided into multiple objects, each containing an individual hierarchy of IEs. While it is possible to group all IEs into a single object, subdividing them into separate IH structures offers several advantages. This approach enhances modularity, making managing, maintaining, and updating specific components easier without affecting the entire system. It also improves scalability, adding new elements or hierarchies seamlessly. Furthermore, a well-structured division aids in clarity and reusability, as distinct modules can be used across different applications or configurations. This design choice ensures a more organized and efficient architecture for the sorting station's operations.

- **IH-AssetInformation:** It encompasses all the IEs representing the assets within the sorting station. Each asset is meticulously detailed with information such as AssetID, AssetName, Connections, FailureRate, SecurityLevel, or a List of Vulnerabilities. These attributes provide a comprehensive view of each asset, facilitating efficient management and risk assessment.

All relevant details are stored as IE instances specific to their respective assets, ensuring structured and easily retrievable data. This hierarchical organization enables seamless integration, analysis, and updates, supporting decision-making processes.

Fig. 6.5 illustrates how asset information is stored and interconnected, showcasing the relationships and dependencies between various assets within the sorting station. This interconnected structure helps identify critical points of failure, understand security vulnerabilities, and optimize overall system performance.

- **IH-ProcedureOfSortingStation:** This object encapsulates the entire operational procedure of the sorting station, structured hierarchically to include each step and its associated elements. It defines the workflow from the initial placement of the workpiece on the conveyor to its sorting and retrieval. The object includes IEs that describes the normal operation of the sorting station.

The completion of one process may depend on the previous process and/or asset availability. This is denoted by the connection between the IH of AssetInformation and IH of procedure. This IH object ensures a clear understanding of the procedural

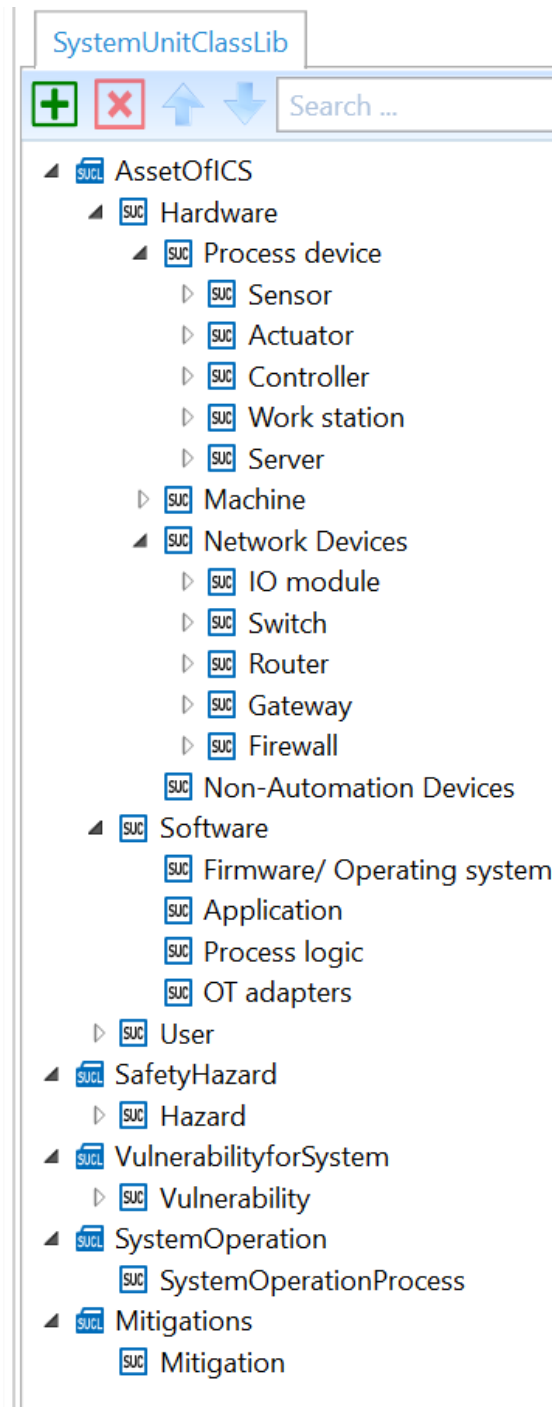


Figure 6.3: System unit classes of Assets in ICS

flow, providing a foundation for system analysis, troubleshooting, and optimization.




































- ▲  AssetInformation
  - ▷  Programmable logic Device
  - ▷  Optical Sensor at Start of conveyor belt
  - ▷  Optical Sensor at End of conveyor belt
  - ▷  Motor for movement
  - ▷  Forked light barrier sensor for detection module
  - ▷  Solenoid deflector for red work piece
  - ▷  Solenoid deflector for Black workpiece
  - ▷  Light barrier sensor red
  - ▷  Light barrier sensor Black
  - ▷  Cobot Interface
  - ▷  Cobot
  - ▷  Maintenance
  - ▷  Operator
  - ▷  Emergency stop
  - ▷  Compressor
- ▲  ProcedureOfSortingStation
  - ▷  Process01: The workpiece is detected at the start of the Sorting Inline station's conveyor
  - ▷  Process02: Workpiece in downstream station
  - ▷  Process03: The detection module determines the color of the workpiece
  - ▷  Process04: Deflector for the defined slide is advanced
  - ▷  Process05: Workpiece ejected
  - ▷  Process06: Deflector for the defined slide is retracted
  - ▷  Process07: Workpiece is detected at end of conveyor
  - ▷  Process08: Workpiece picked up from end of conveyor belt
  - ▷  SPC: Sorting Station Process Complete
- ▲  SafetyHazardInstance
  - ▷  Operator Injured
- ▲  Mitigations
  - ▷  Multi-Factor Authentication (MFA)
  - ▷  Regular Patching and Vulnerability Management
  - ▷  Rate limiting to the device
  - ▷  Upgrade PLC firmware to 3.0.1
  - ▷  Restrict physical access to affected devices
  - ▷  improve Training and procedure follow monitoring

Figure 6.4: Instance hierarchy of Use Case

Organizing the procedure into detailed IEs facilitates adaptability and scalability, allowing for seamless updates or modifications as operational requirements evolve.

- IH-SafetyHazardInstance: It represents individual safety hazards identified within

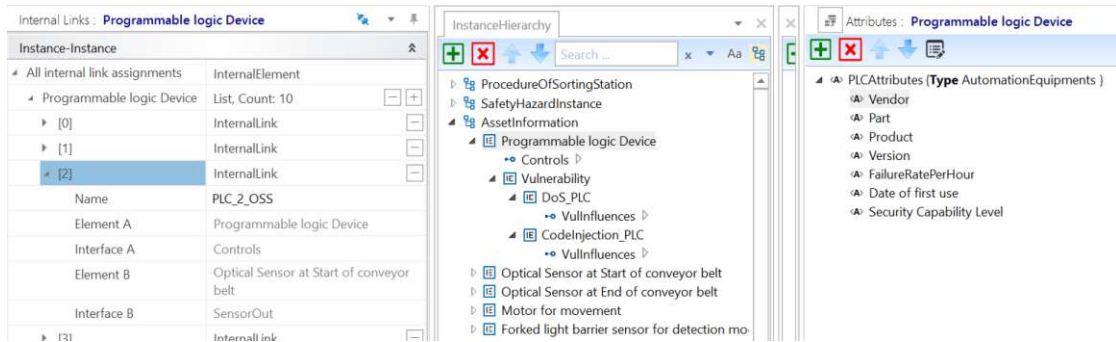


Figure 6.5: Instance hierarchy, Internal Links, and Attributes example of Asset (PLC)

the sorting station. Each instance is structured to include IEs that provides detailed information about the hazard, such as HazardID, HazardDescription, TriggerConditions, AffectedComponents, LikelihoodOfOccurrence, and potential consequences. The IEs allows for a comprehensive understanding of each hazard, enabling effective prioritization and mitigation planning.

- **IH-Mitigations:** It details the measures implemented to mitigate identified hazards and enhance the safety and security of the sorting station. Each mitigation is represented as an IE and includes attributes such as MitigationID, TargetHazard, MitigationDescription, ImplementationDetails, EffectivenessRating, and Status (e.g., planned, in-progress, completed). Mitigations are linked directly to their corresponding hazard instances to provide traceability and ensure alignment with risk management strategies. Additionally, this object may document the responsible personnel, timelines, and resources required for each mitigation action, supporting streamlined implementation and monitoring.

### Role Class Library of the Sorting Station

The classes, the modules, and individual components belong to are defined by the SUCL. The RCL for this model is called *Requirement* for the role of the component. The requirements are: *Process, Safety, Security, and Communication* as shown in Fig. 6.6.

The *process requirement* pertains to the component's role in executing and managing the core industrial processes. It involves monitoring, controlling, and regulating physical or chemical processes within the system. Components associated with the Process requirement maintain desired operational parameters, coordinate actions between various subsystems, and ensure smooth and efficient process execution.

The *safety requirement* focuses on components that ensure the safety of the ICS and the personnel involved. These components are designed to detect and prevent hazardous situations, mitigate risks, and safeguard against potential accidents or incidents. They may include safety interlocks, emergency shutdown systems, protective barriers, or redundant control mechanisms to enhance system resilience and minimize possible harm.

The *security requirement* addresses the protection of the ICS from unauthorized access, malicious attacks, and potential cybersecurity threats. Components associated with the Security requirement implement authentication, access control, encryption, intrusion detection, and vulnerability management. These measures safeguard critical infrastructure, data integrity, and confidentiality, helping to prevent unauthorized modifications, data breaches, or disruptions to system operations.

The *communication requirement* involves components facilitating communication and information exchange within the ICS. These components enable seamless and reliable data transmission between various subsystems, devices, and human-machine interfaces. They may utilize protocols, networks, or interfaces to support real-time monitoring, control commands, data logging, and reporting. Reliable communication ensures efficient coordination and integration of system components, enabling smooth operations and effective decision-making.

By defining and understanding these requirements, stakeholders in ICSs can effectively design, implement, and maintain components based on their intended roles and responsibilities. This classification framework promotes a systematic approach to system development, ensuring that each element aligns with specific requirements and fulfills its designated functions. Ultimately, this contributes to the overall performance, safety, security, and reliability of the ICS.

Other than these, one can have additional requirements like reliability and availability to ensure uninterrupted operation, scalability to accommodate system changes, interoperability for seamless integration of components, real-time performance for timely control, robustness to withstand harsh environments, regulatory compliance to meet standards and guidelines, data integrity and backup to safeguard critical information, user-friendly interfaces for efficient human-machine interaction, comprehensive training and documentation, and disaster recovery and business continuity planning. These additional requirements can be sub-classes of the core requirements mentioned above. Addressing these requirements alongside the core ones is essential for successfully implementing, operating, and maintaining an ICS, ensuring optimal performance, reliability, safety, and security.



Figure 6.6: Role Class Library for Sorting Station

### EI and ILs of the Sorting Station

In the context of the AML model, there are two important concepts: ILs and EIs. These elements are crucial in defining the connectivity and interactions within the AML model and its integration with external systems.

An IL refers to the connection between two elements or components within the AML model itself. It establishes relationships and dependencies between different parts of the model, allowing them to exchange data or interact with each other. ILs represent the logical connections and associations between modules, devices, and other entities within the AML model. The ILs are shown in Fig. 6.7 as the lines between the components.

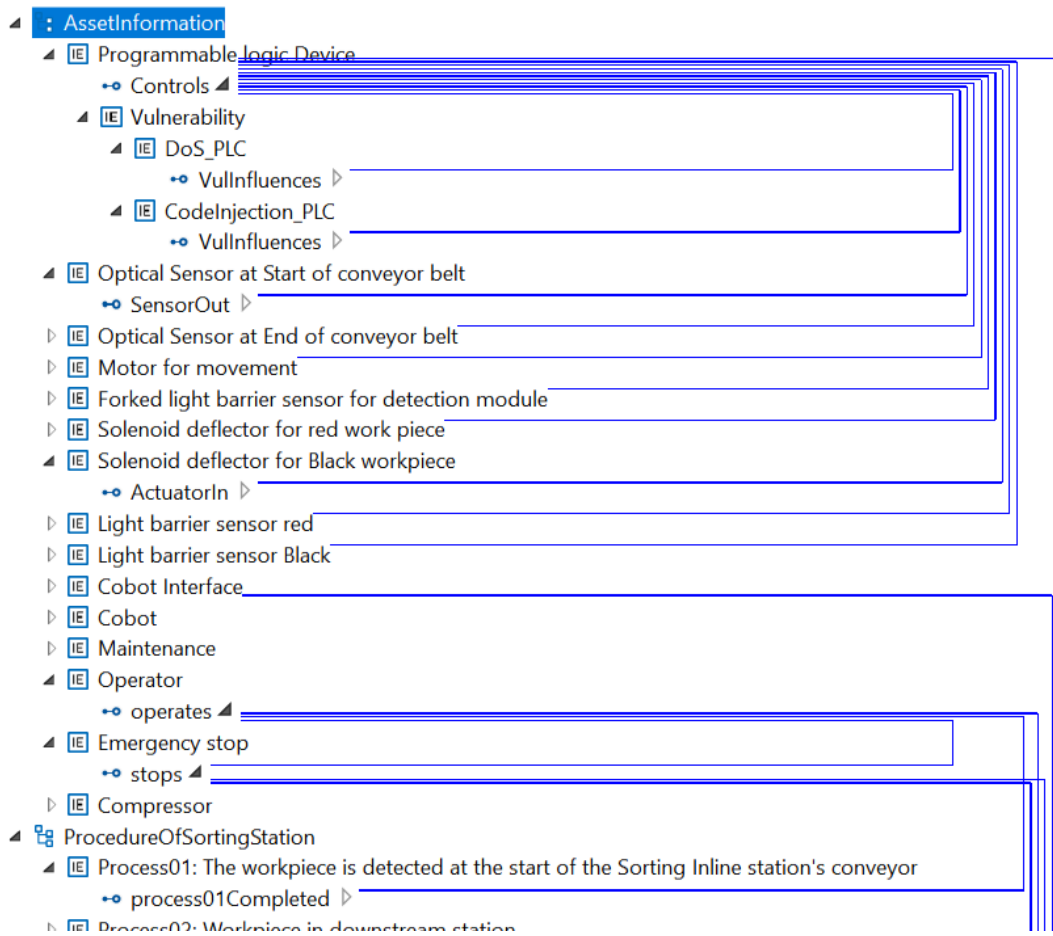


Figure 6.7: Expansion of Instance hierarchy with ILs

On the other hand, an EI enables the IL connection. It allows communication and data exchange between the AML model and other software or hardware systems outside its boundaries. EIs are essential for integrating the AML model with other automation systems, control systems, user interfaces, or external devices.

## Interface Class Library for Sorting Station

The Interface library in the AML model provides various types of connections to facilitate different interactions, as shown in Fig. 6.8. These connection types include *network-based*, *logic-based*, *user-based*, *HazardRef*, and *VulnerabilityRef* connections. However, one can only use one connection reference for all connections.

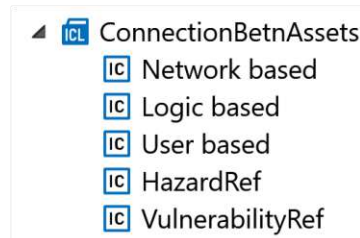


Figure 6.8: Interface class for the sorting station

The model shows two connection types: *Network-based* and *Logic-based*. The actual network connection is depicted in *Network-based*, where the PLC controls the OSE. Still, the exact connection is from PLC → switch → IO-Link gateway → mini IO → the actual sensor. *PLC\_Switch\_NWLink* goes from Element A (PLC) to Element B (Switch). In *Logic-based*, the connection is based on the actual program logic connection. According to the automation pyramid, the components' connection is depicted here. *PLC\_OSE\_LogicLink* goes from Element A (PLC) to Element B (optical sensor). Similarly, the *User-based* ILS depict the work function of the humans (Operator and maintenance) involved in the operation.

## Attribute Class Library of Sorting Station

Attributes provide additional information about the elements, such as their physical or logical properties, operational parameters, or behavior. The attribute class is a data structure containing various properties and values. It specifies the attribute type, its name, data type, and other relevant information. The attribute class can be customized to match the specific requirements of the automation system. Using attribute classes, AML enables interoperability and data exchange between engineering tools, automation systems, and domains, promoting seamless integration and collaboration in automation engineering.

ICS assets rely on various attributes for effective management and maintenance. The *Vendor* attribute provides information about the manufacturer or supplier of the asset, enabling access to technical support, warranty claims, and documentation. It establishes direct communication, ensuring reliable assistance and timely updates for optimal asset performance and security. The *Part* attribute is a unique identifier, facilitating accurate replacement and maintenance activities while aiding inventory management to track and control spare parts efficiently.

Additionally, the *Product* attribute plays a crucial role in asset identification and referencing within the organization's inventory system. It streamlines asset tracking, maintenance planning, and documentation, distinguishing between different variations or versions of assets and ensuring compliance with specifications during repairs or upgrades. Assessing reliability is vital, and the *Failure Rate* attribute quantifies the probability or frequency of asset failure. This information enables prioritized maintenance, effective resource allocation, and proactive strategies to minimize unplanned downtime.

For software-driven assets, the *Version* attribute denotes the specific release or iteration of the installed software or firmware. It is essential for updates, security patches, bug fixes, and compatibility with other components or systems. Monitoring and managing the *Version* attribute ensures assets are running the latest software, benefiting from performance enhancements, critical security fixes, and improved compatibility. By leveraging these attributes, organizations can optimize asset management, minimize disruptions, and enhance overall operational efficiency.

In addition to these attributes, several other attributes can be helpful for ICS assets. For example, the *Location* attribute specifies the physical or logical placement of the asset within the industrial environment. The *Criticality* attribute assesses the asset's importance to the overall system or process. The *Maintenance Schedule* attribute outlines recommended maintenance activities and intervals. The *Service History* attribute captures a log of the asset's maintenance activities, repairs, and interventions conducted. The *Compliance* attribute indicates whether the asset meets specific regulatory or industry standards.

Risk assessment can be effectively conducted using AML by leveraging its structured and standardized framework to represent the system's architecture, the inventories of assets, and the associated hazards and vulnerabilities. The AML model provides a unified platform to integrate various data sources, including expert insights, system designs, asset specifications, process workflows, and database records. This harmonized data structure enables the application of risk assessment methodologies to analyze potential risks, evaluate their impact, and devise mitigation strategies.

Integrating AML with risk assessment processes ensures a comprehensive approach to managing risks in complex systems. By combining detailed engineering data, run-time information, and historical records, AML facilitates the identification of interdependencies and potential failure points. This enables risk managers to simulate scenarios, predict outcomes, and prioritize mitigation measures based on their significance. Furthermore, the graphical and hierarchical nature of AML supports clear visualization and traceability of risks, ensuring alignment with organizational safety and security goals.

### 6.1.2 Limitations

Despite its strengths, AML has limitations when applied to risk assessment:

- Complexity of Modeling: Building an exhaustive AML model for a large and

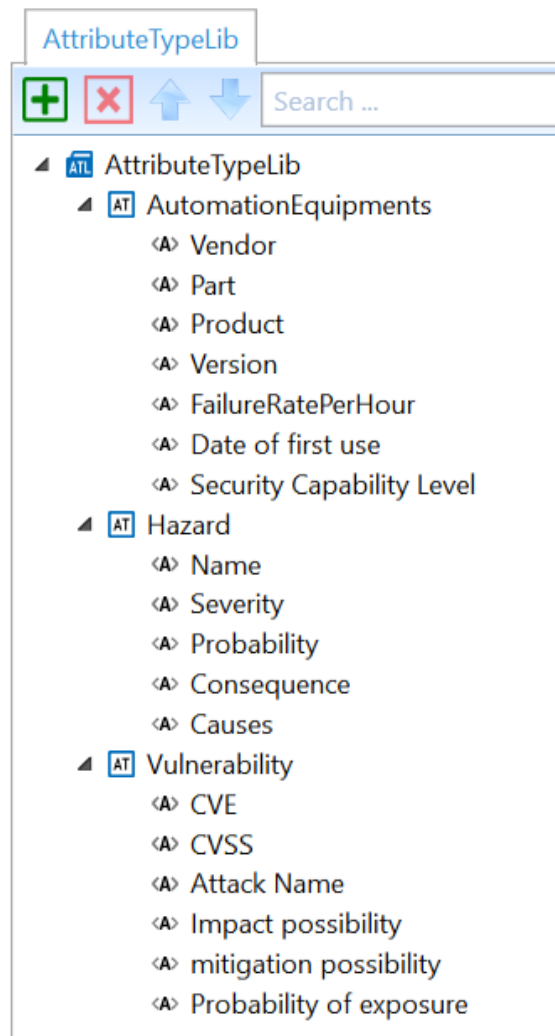


Figure 6.9: Attributes for Assets of ICS

intricate system can be time-consuming and resource-intensive. This complexity may hinder rapid assessments in dynamic environments.

- **Dependence on Data Quality:** The effectiveness of AML in risk assessment is highly dependent on the accuracy and completeness of the input data. Incomplete or outdated information can lead to inaccurate risk assessments.
- **Limited Real-Time Integration:** While AML excels in static modeling, it may face challenges in real-time integration with operational data, limiting its application for dynamic risk monitoring.
- **Expertise requirement:** Using AML requires expertise and familiarity with its standards and tools. Organizations with limited technical expertise may encounter

a steep learning curve.

- Tooling and Interoperability: The adoption of AML often depends on compatible tools and software. The academic version has limited modeling possibilities.
- Versionability Concerns: For highly scalable systems with rapidly evolving architectures, maintaining and updating AML models can become a bottleneck, reducing their practicality.

## 6.2 Asset Administration Shell

The AAS <sup>2</sup> is a framework to create digital representations of physical assets in the context of Industry 4.0. AAS is intended to provide a standardized way to represent assets, capabilities, and interactions with other assets in a connected ecosystem.

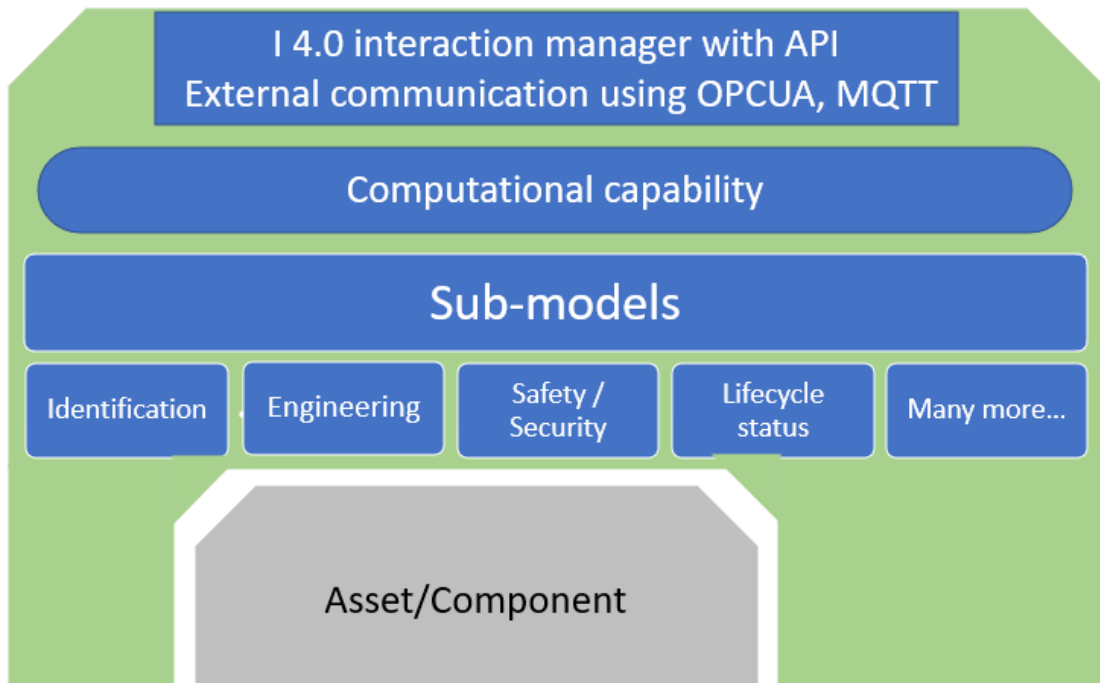


Figure 6.10: AAS structured implementation [14]

The AAS framework can be used to create digital representations of various aspects of an industrial plant, enabling better management, monitoring, and optimization of industrial operations. The AAS can be used for various purposes, including the following:

- Asset Inventory: AAS can create a digital inventory of assets in an industrial plant, including their properties, such as location, status, and maintenance history.

<sup>2</sup><https://www.plattform-i40.de/IP/Redaktion/EN/Standardartikel/specification-administrationshell.html>

- **System Architecture:** AAS can be used to create a digital representation of the system architecture of an industrial plant, including the components, connections, and interfaces between them.
- **Process Definition:** AAS can be used to create a digital representation of processes in an industrial plant, including their inputs, outputs, and dependencies.
- **Databases:** AAS can be used to create a digital representation of databases in an industrial plant, including their structure, schema, and relationships between tables.
- **Safety and Security Policies and Requirements:** AAS can be used to create a digital representation of an industrial plant's safety and security policies and requirements, including risks, hazards, and mitigation measures.

### 6.2.1 AAS Structured Representation

The structure of AAS is defined by a technology-independent meta-model, which is complemented by multiple technology-specific serialization formats, including XML, JSON, and OPC UA [140]. AAS instances can be interconnected using the external reference feature, enabling the representation of hierarchical system architectures.

Each AAS encapsulates asset-related information in the form of submodels, which store data such as technical specifications, operational parameters, maintenance records, and communication interfaces. Within a submodel, the Submodel Elements Collection (SMC) further organizes related attributes, for instance, grouping all parameters associated with operational data. Properties within the AAS framework function similarly to the characteristics in AML, representing individual values or parameters of the asset.

Figure 6.11 illustrates the model-based structure and relationships between various AAS components.

AAS facilitates seamless interoperability between diverse systems and components by standardizing data exchange formats, ensuring smooth integration across tools and platforms <sup>3</sup>.

### 6.2.2 AAS model creation for Use Case

The AAS representation handles information related to an asset. An asset can be a tangible (physical sensors, controllers, etc.) or intangible (software, process, etc.) component of the system. One needs to implement multiple AAS models separately to address various components. However, the connection between components and dependencies is addressed by the *reference* method.

Referencing other AAS models within a primary AAS model allows the creation of a hierarchical digital representation of complex systems. This approach mirrors the

<sup>3</sup><https://tinyurl.com/5yf8nwzb>

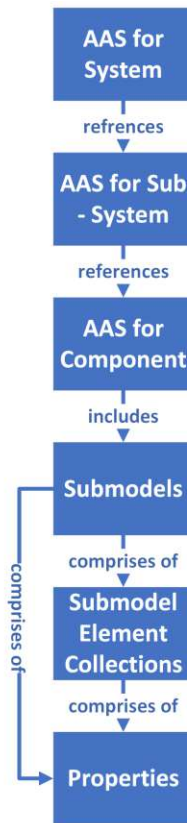


Figure 6.11: AAS model

physical structure of systems by allowing an overarching system-level AAS to connect with multiple subsystem and component-level AAS models. Each AAS serves as a digital twin, encapsulating detailed information about the corresponding physical asset.

The external reference feature facilitates the connection between the AAS models. Through this mechanism, a system-level AAS can point to other AAS instances using globally unique identifiers such as Internationalized Resource Identifiers (IRIs) or Uniform Resource Names (URNs). These identifiers ensure that each AAS is uniquely identifiable and can be referenced consistently across platforms. Within the AAS structure, submodels play a crucial role in organizing the information, and the reference elements within these submodels are used to point to external AAS models. The ability to reference other AAS models improves scalability, modularity, and interoperability [20].

Fig. 6.12 provides a glimpse of all the AAS of the components, the identified hazards, and the operational procedures. After the assessment, mitigations can also be identified and connected to the particular component or the overall system through mitigation AAS.

Package	C:\Users\Admin\Desktop\all folders\AAS_RA\aaax files\Sorting_station_NEW_AASX_File_v2.aasx
Env	Environment
Env	AdministrationShells
AAS	PLC [https://safeseclab.com/MPS/SS/aas/PLC] of [https://safeseclab.com/MPS/SS/PLC, Instance]
AAS	OSS [https://safeseclab.com/MPS/SS/aas/OSS] of [https://safeseclab.com/MPS/SS/OSS, Instance]
AAS	OSE [https://safeseclab.com/MPS/SS/aas/OSE] of [https://safeseclab.com/MPS/SS/OSE, Instance]
AAS	MTR [https://safeseclab.com/MPS/SS/aas/MTR] of [https://safeseclab.com/MPS/SS/MTR, Instance]
AAS	FLBS [https://safeseclab.com/MPS/SS/aas/FLBS] of [https://safeseclab.com/MPS/SS/FLBS, Instance]
AAS	SD_RS [https://safeseclab.com/MPS/SS/aas/SDR] of [https://safeseclab.com/MPS/SS/SDR, Instance]
AAS	SD_BS [https://safeseclab.com/MPS/SS/aas/SDB] of [https://safeseclab.com/MPS/SS/SDB, Instance]
AAS	LBSR [https://safeseclab.com/MPS/SS/aas/LBSR] of [https://safeseclab.com/MPS/SS/LBSR, Instance]
AAS	LBSB [https://safeseclab.com/MPS/SS/aas/LBSB] of [https://safeseclab.com/MPS/SS/LBSB, Instance]
AAS	Comp [https://safeseclab.com/MPS/aas/COMP] of [https://safeseclab.com/MPS/COMP, Instance]
AAS	OPR [https://safeseclab.com/MPS/SS/aas/OPR] of [https://safeseclab.com/MPS/SS/OPR, Instance]
AAS	MNC [https://safeseclab.com/MPS/SS/aas/MNC] of [https://safeseclab.com/MPS/SS/MNC, Instance]
AAS	CBTI [https://safeseclab.com/MPS/SS/aas/CCT] of [https://safeseclab.com/MPS/SS/CCT, Instance]
AAS	CBT [https://safeseclab.com/MPS/SS/aas/CBT] of [https://safeseclab.com/MPS/SS/CBT, Instance]
AAS	ES [https://safeseclab.com/MPS/SS/aas/ES] of [https://safeseclab.com/MPS/SS/ES, Instance]
AAS	SafetyHazard [https://safeseclab.com/MPS/SS/aas/SafetyHazard] of [https://safeseclab.com/MPS/SS/Hazard, Instance]
AAS	Procedure [https://safeseclab.com/MPS/SS/aas/Procedure] of [https://safeseclab.com/MPS/SS/Procedure, Instance]
AAS	Mitigation [https://safeseclab.com/PS/SS/aas/Mitigation] of [https://safeseclab.com/MPS/SS/Mitigation, Instance]

Figure 6.12: All AASs for the use case Sorting Station.

AAS	PLC [https://safeseclab.com/MPS/SS/aas/PLC] of [https://safeseclab.com/MPS/SS/PLC, Instance]
Asset	AssetInformation https://safeseclab.com/MPS/SS/PLC
SM	Reliability [https://safeseclab.com/MPS/SS/aas/PLC/submodel_rel]
SM	Nameplate [https://safeseclab.com/MPS/SS/aas/PLC/submodel_nameplate]
SM	Vulnerability [https://safeseclab.com/MPS/SS/aas/PLC/submodel_vul]
SM	Connection [https://safeseclab.com/MPS/SS/aas/PLC/submodel_conn]
SM	Security_Data [https://safeseclab.com/MPS/SS/aas/PLC/submodel_secure]

Figure 6.13: An example of PLC AAS along with submodels.



Figure 6.14: All submodels from PLC AASs for Sorting Station.

### 6.2.3 Limitations

Despite its strengths, AAS has limitations when applied to risk assessment and industrial applications:

- **Complexity of Modeling:** Developing a comprehensive AAS model for large-scale industrial systems can be complex and time-intensive. Defining appropriate submodels for diverse assets requires significant effort.
- **Dependence on Data Quality:** The effectiveness of AAS relies on the accuracy and completeness of input data. Inconsistent or outdated data can lead to unreliable asset representations and poor decision-making.
- **Limited Real-Time Integration:** While AAS enables structured asset information, it has limited support for real-time data exchange and high-frequency control operations, making it less suitable for dynamic process control.
- **Expertise Requirement:** Implementing AAS requires knowledge of industry standards, ontologies, and interoperability frameworks. Organizations with limited technical expertise may struggle with implementation.

- **Tooling and Interoperability:** The adoption of AAS depends on available software tools and compatible digital twin platforms. Standardized submodels for industrial use cases are still evolving.
- **Scalability:** Managing and updating AAS models for large-scale, complex industrial environments can become challenging, especially when assets undergo frequent changes or require real-time synchronization.

## 6.3 Ontology

Ontology structures data and relationships to enhance intelligent systems and knowledge management. The Resource Description Framework (RDF) models linked data using classes and properties, enabling structured and reusable information exchange. The Web Ontology Language (OWL) extends RDF with a richer vocabulary of classes, properties, and relations, facilitating semantic data integration across domains. OWL represents information as triples: Subject (entity), Predicate (relationship), and Object (associated value or entity).

The design of an ontology is inherently constrained by its competency questions, which define the scope and granularity of the modeled knowledge. In this context, the ontology must effectively address the following key questions:

- What are the assets in the system, and what is their type?  
Ensuring a structured classification of assets, attributes, and roles within the system.
- How are the assets connected for system operation?  
Capturing relationships, dependencies, and interactions essential for operational functionality.
- What are the identified hazards for processes and users, and what vulnerabilities exist in OT services?  
Enabling risk assessment by modeling threats, vulnerabilities, and their impact on system security and safety.

### 6.3.1 Ontology Structured Representation

An ontology formally defines concepts and relationships within a system, whether for humans, machines, or agents [119]. Ontology-based approaches organize these elements to minimize ambiguity, enhance knowledge representation, and ensure consistency and accuracy. They are widely applied in fields such as artificial intelligence, knowledge management, the semantic web, and information systems [12].

An ontology is a structured framework for formally representing knowledge, defining a domain's concepts, entities, relationships, and properties [12, 108, 119]. By leveraging an

ontology-based approach, a domain can be systematically modeled, capturing its essential characteristics in a structured manner. This model is a foundation for analyzing and optimizing the system, enabling a better understanding and reasoning about its behavior.

A formal ontology consists of key components defining a precise and structured domain representation. These include Classes, which categorize similar entities; Relations or Properties, which establish connections between entities; Attributes or Datatypes, which define entity characteristics; and Individuals or Instances, which represent specific examples of entities within a class [12]. Fig. 6.15 represents all the entities in the ontology diagram and their relationships.

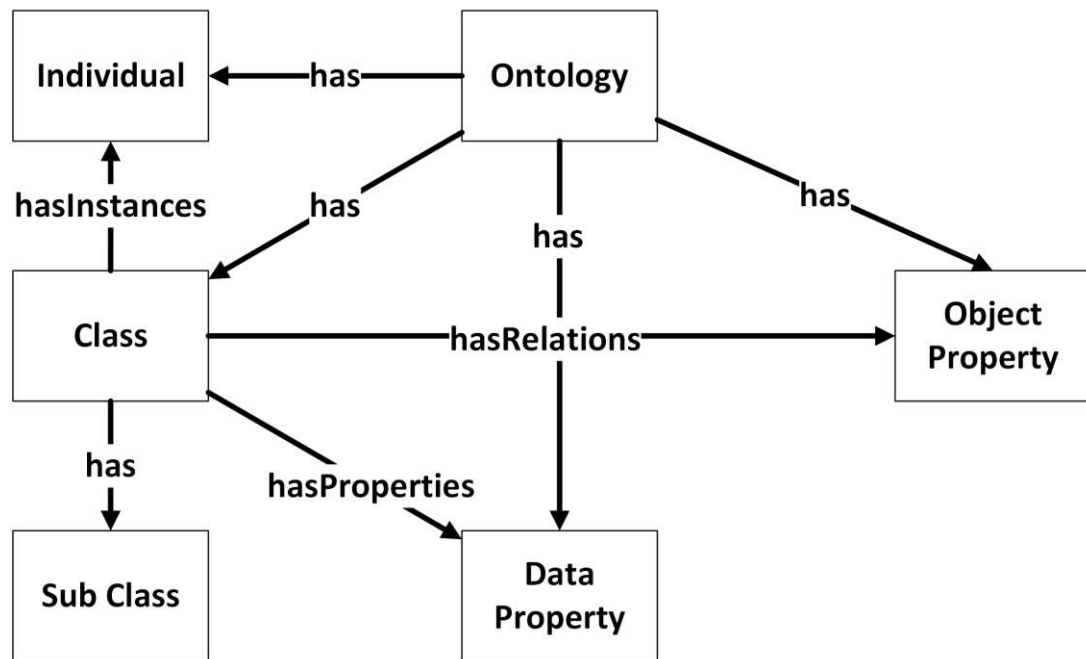


Figure 6.15: Ontology entities representation

### 6.3.2 Ontology Model creation for Use Case

The model creation for the ontology is done using Protégé<sup>4</sup>. It is a software tool that is available freely for creating and using ontologies. It is an open-source ontology editor and framework for building intelligent systems. It supports W3C standards, plug-ins, and a large community of users and developers. WebProtégé<sup>5</sup> is an ontology development environment for the Web that makes it easy to create, upload, modify, and share ontologies for collaborative viewing and editing.

<sup>4</sup><https://protege.stanford.edu/>

<sup>5</sup><https://webprotege.stanford.edu/>

Fig. 6.16 presents an ontology that defines various classes and their interrelations within a specific use case. The main classes in this ontology include SystemAsset, SystemOperationProcess, Mitigation, Threat, Vulnerability, and Hazard. These classes represent essential elements and their interactions in a system, particularly in risk management and security assessment.

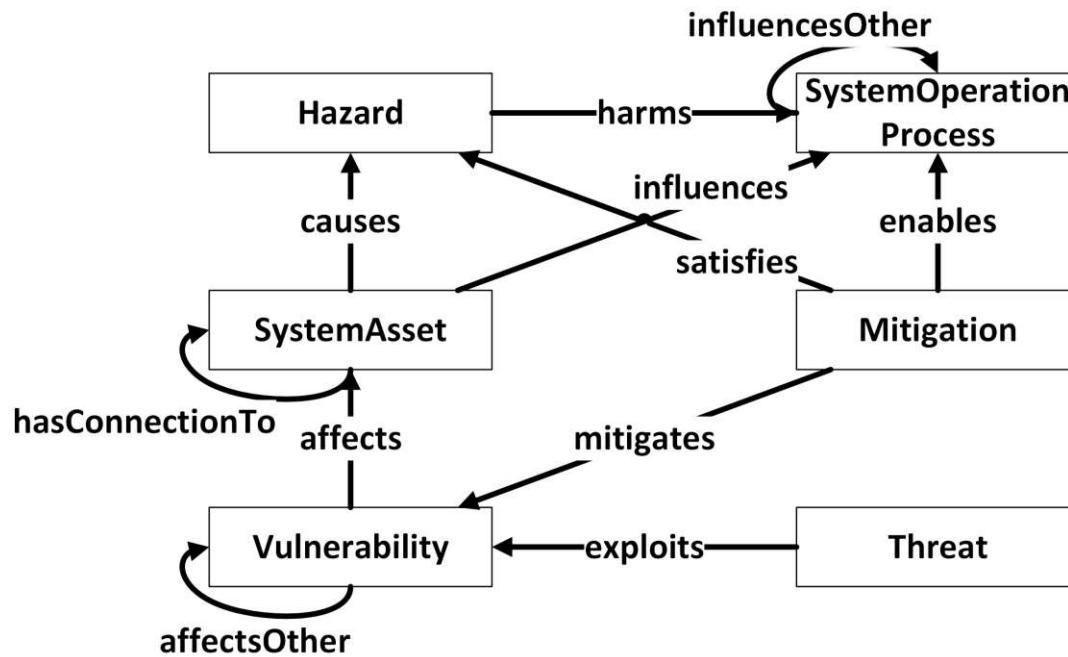


Figure 6.16: Ontology classes and their relations for Use Case

SystemAsset represents tangible or intangible components within a system. It can be connected to other SystemAssets, denoting dependencies or physical/logical connections. Furthermore, a SystemAsset may cause a Hazard, implying that its malfunction or misconfiguration could lead to a risk situation. Additionally, a SystemAsset can influence a SystemOperationProcess, indicating that assets play a role in the execution or effectiveness of operational processes.

SystemOperationProcess represents actions or workflows carried out within the system. One SystemOperationProcess can influence another, meaning that processes are interdependent and changes in one can affect others. Additionally, external factors like SystemAssets can influence how these processes function.

Mitigation plays a crucial role in reducing risks. A Mitigation can satisfy a Hazard, meaning it addresses or fulfills the need to counteract a potential risk. It can also mitigate a Vulnerability, reducing the likelihood of it being exploited. Moreover, a mitigation may enable a SystemOperationProcess, indicating that implementing risk-mitigation strategies can facilitate or support operational activities.

Threats and Vulnerabilities are critical risk factors in ontology. A Threat exploits a Vulnerability, highlighting how weaknesses in the system can be taken advantage of by potential risks. Additionally, a Vulnerability can affect another Vulnerability, showing how weaknesses might propagate within a system. Fig. 6.17 shows the whole ontology representation.

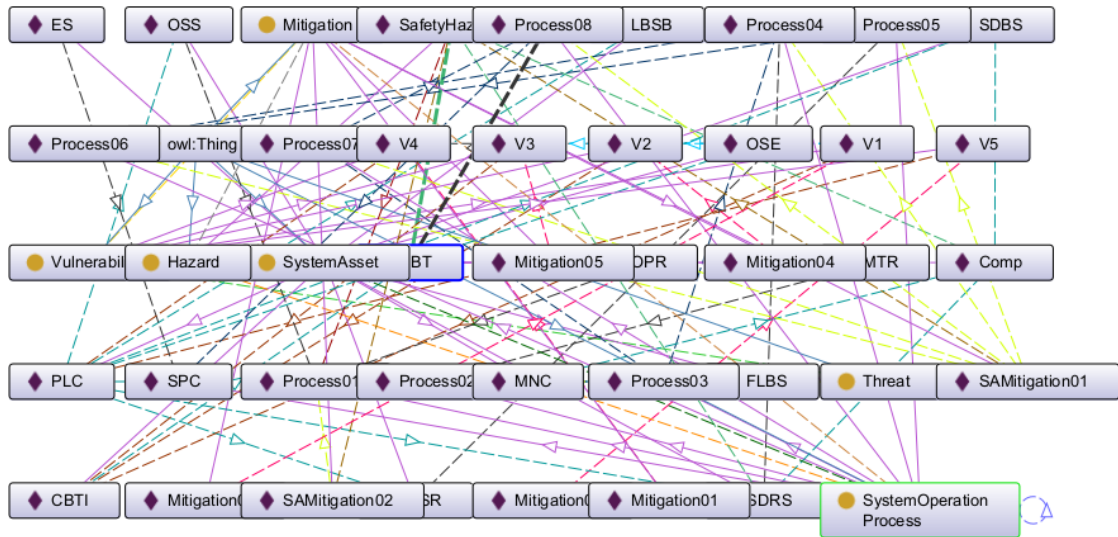


Figure 6.17: Complete Ontology graphical representation using OntoGraph

Fig. 6.18 illustrates how the ontology is represented in Protégé, a widely used ontology editor. It provides a structured view of the ontology through different aspects, including classes, properties, and instances.

The reasoner in Protégé is a tool used to automatically infer logical conclusions from the ontology's axioms, relationships, and constraints. It ensures that the ontology is logically consistent and helps classify concepts based on their definitions. In Protégé, reasoners perform various tasks, such as checking for inconsistencies, inferring new relationships, and classifying the ontology structure. Fig. 6.19 represents the result of the reasoner used for inconsistency checking on the ontology model.

The reasoner used is HermiT 1.4.3.456, a well-known OWL 2 reasoner. HermiT is efficient and widely used for handling complex ontologies with expressive logic. It is used for ontology classification, consistency checking, and inference generation. It ensures that the ontology is logically sound and helps uncover implicit relationships, making it an essential tool for ontology development and validation.

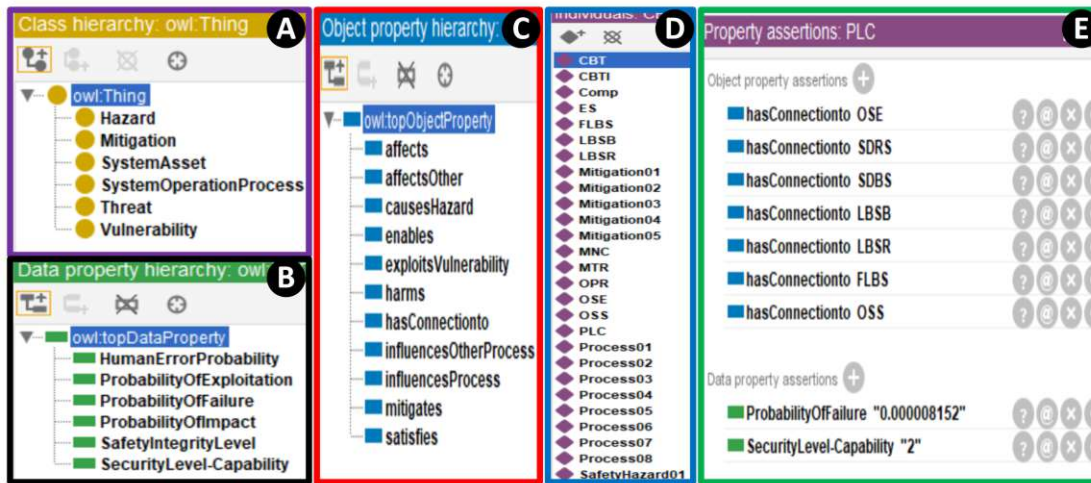


Figure 6.18: Representation of the Ontology in the Protégé in terms of (A) Classes, (B) Data properties, (C) Object properties, (D) Individual instances, and (E) Instance relation representation with other instances using object properties and data properties

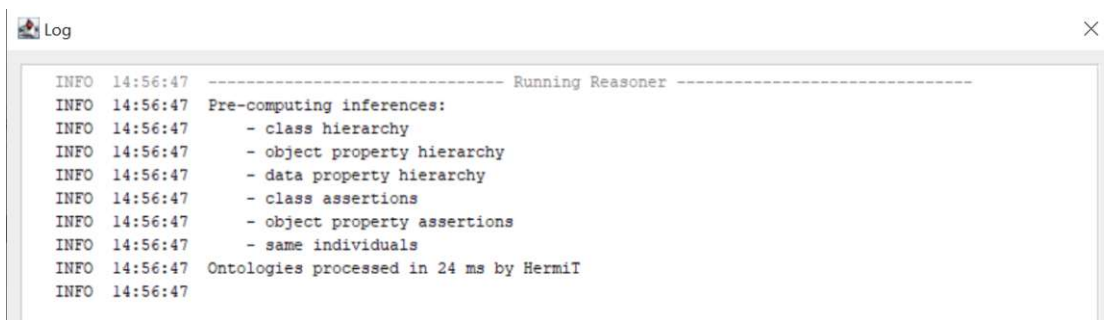


Figure 6.19: Reasoner result of the Ontology

### 6.3.3 Limitation

Ontology development and tools like Protégé are essential for structured knowledge representation, but they have complexity, reasoning performance, scalability, and integration challenges. While Protégé is a powerful tool, its performance, visualization, and collaboration limitations require careful consideration when working with large and evolving ontologies.

- **Complexity and Development Effort:** Designing a well-structured ontology requires significant time and expertise.
- **Scalability Issues:** Large ontologies can lead to performance challenges in querying and reasoning.

- **Reasoning Performance:** Ontology reasoners may become slow when processing extensive datasets.
- **Limited Representation of Uncertainty:** Ontologies struggle to model probabilistic or uncertain knowledge.
- **Maintenance and Evolution:** Updating ontologies to reflect new knowledge is labor-intensive and may introduce inconsistencies.

## 6.4 Information Utilization

In this chapter, we examined the data sources required for risk assessment in the sorting station application of a ICS. We analyzed how they are represented using different information models. Each approach offers unique advantages: AML provides a structured format for engineering data exchange, AAS enables standardized digital representations of assets, and Ontology facilitates semantic relationships and reasoning capabilities.

Understanding how each technology structures and processes risk-related data is crucial for practical risk assessment. The structured information collected in this phase serves as the foundation for the next step, where we analyze risks by identifying relationships between system components, vulnerabilities, threats, and mitigations.

Criteria	AML	AAS	Ontology
<b>Purpose</b>	Standard for exchanging engineering data in industrial automation	Digital representation of industrial assets with structured submodels	Formal representation of domain knowledge with reasoning capabilities
<b>System Architecture Representation</b>	Models system components using CAEX, IEC 61131-3, and COLLADA standards	Represents system architecture through submodels but lacks standardized physical modeling	Can formally describe system components and their relationships with logical inference support
<b>Asset-Related Information</b>	Defines assets with hierarchical structures, metadata, and engineering properties	Captures asset details in submodels (e.g., Identification, Technical Data, Maintenance)	Provides semantic descriptions of assets, enabling advanced queries and reasoning

Continued on next page

(Continued from previous page)

Criteria	AML	AAS	Ontology
<b>Asset Connections</b>	Uses topology structures to define asset interconnections but is limited to predefined schemas	Represents asset connections using references within submodels but lacks detailed semantic relationships	Defines complex relationships using object properties, enabling advanced reasoning on asset interactions
<b>Identified Hazards</b>	Can store hazard-related metadata, but no built-in support for risk assessment logic	Can document hazards in dedicated submodels but lacks reasoning capabilities	Represents hazards as concepts, linking them to assets, vulnerabilities, and mitigation measures with logical inference support
<b>Identified Vulnerabilities</b>	Can store vulnerability data within system components but lacks built-in risk analysis	Can structure vulnerability information within specific submodels but lacks automated assessment	Models vulnerabilities semantically and connects them to threats, risks, and mitigations for automated analysis
<b>System Process Representation</b>	Describes industrial processes using automation standards but is limited in flexibility	Models process-related information in submodels but lacks reasoning and inference capabilities	Defines process relationships, dependencies, and constraints with logic-based reasoning
<b>Support for Risk Assessment</b>	Limited support for structured risk assessment; requires external tools for analysis	Can document risk-related data in submodels but lacks inference capabilities	Provides a formal risk assessment framework by linking hazards, threats, vulnerabilities, and mitigation strategies
<b>Continued on next page</b>			

(Continued from previous page)

Criteria	AML	AAS	Ontology
<b>Integration with Other Systems</b>	Strong integration with industrial automation tools but limited semantic interoperability	Designed for I4.0 interoperability with OPC UA and digital twins	Easily integrates with semantic web technologies and knowledge graphs for interoperability
<b>Reasoning and Decision Support</b>	No built-in reasoning or inference mechanisms	Lacks reasoning; mainly serves as a structured data model	Supports automated reasoning, enabling advanced decision-making and risk prediction

Table 6.1: Comparative analysis of information models as the single source of truth for risk assessment

The information provided by each information model is extracted and utilized as the foundation for the risk assessment framework. Serving as a single source of truth, the information model ensures consistency and accuracy in risk analysis. Moreover, adopting information models is essential for enabling automation and facilitating seamless integration and decision-making in ICSs.

# Risk assessment using Bayesian Belief Networks

This chapter discusses the need for a BBN for the safety and security integrated risk assessment. Through various formulae and equations, it also represents the implementation strategy of BBN for the risk assessment.

## 7.1 Single source of truth Model

In comparing AAS, AML, and Ontology, we selected AML for the single source of truth of information used for the risk assessment. AML is a structured and standards-compliant information model. AML facilitates organized development through its foundation on the IEC 62424 standard via the CAEX format, while ensuring data exchange compatibility with IEC 62714. As a proof of concept, AML was preferred due to its superior tooling—most notably the AML Editor—and the availability of comprehensive documentation and literature support [33, 34].

Fig. 7.1 illustrates the conceptual framework for information collection and representation within an AML model, structured around five core elements. The necessary data for risk assessment is sourced from various origins, as depicted on the left side of the diagram. These sources contribute information about vulnerabilities, hazards, assets, and their interrelationships, which are highlighted in blue blocks. This collected data forms the basis for constructing a semantically rich and interconnected AML model. The detailed implementation is provided in Section 6.1.1. However, the visual representation for the implementation is provided in Fig. 7.2.

The developed AML model is designed based on a sorting station use case and is structured following the conceptual methodology illustrated in Fig. 7.1. Only selected and essential

## 7. RISK ASSESSMENT USING BAYESIAN BELIEF NETWORKS

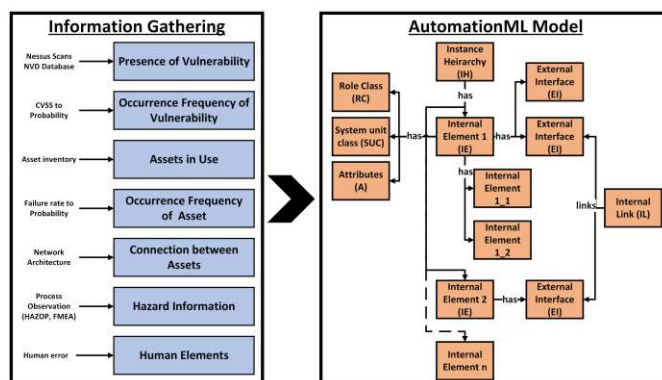


Figure 7.1: Concept of AML model development [19]

components are shown in Fig. 7.2, rather than the complete model representation, to support a clearer understanding of the model's structure and implementation.

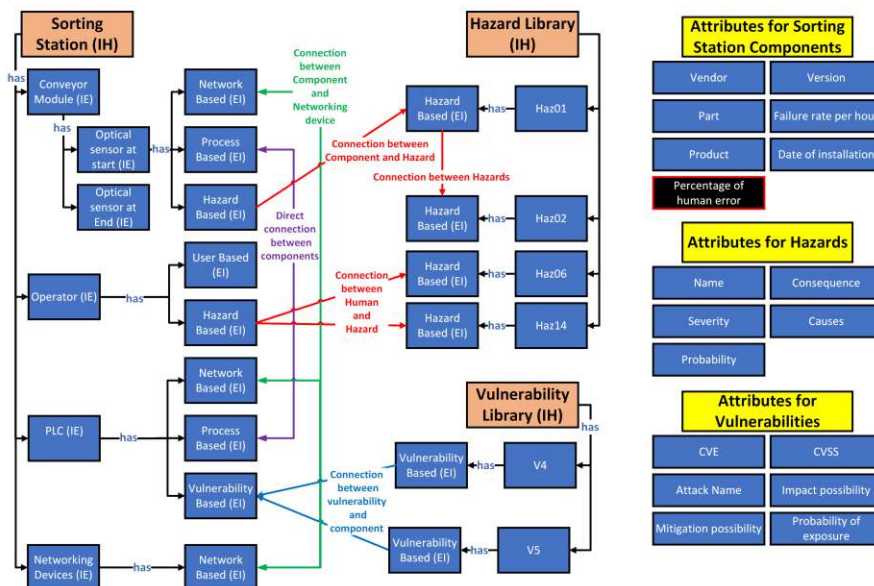


Figure 7.2: AML model implementation

As depicted in Fig. 7.2, it represents the three main IHs: the Sorting Station, which means the main use case; the Hazard Library, which includes risk elements such as those arising from failures or operational errors; and the Vulnerability Library, which catalogs weaknesses inherent in system components. Within the Sorting Station IH, IEs include modules such as the conveyor belt, detection units, and sorting mechanisms, along with sensors like optical detectors positioned at both ends of the conveyor. Additional elements include a PLC and a human operator. The Hazard Library contains IEs such as Haz01, Haz02, Haz06, and Haz14, each representing a specific hazardous condition. In

parallel, the Vulnerability Library consists of elements such as V4 and V5, which denote vulnerabilities in components like the PLC.

Each IE in the model can contain one or more EIs that define its interaction with other elements. These interfaces are categorized into several types, including network-based interfaces (e.g., communication between sensors and the PLC through networking devices), process-based interfaces (e.g., logical connections within control flows), hazard-based interfaces (e.g., operator actions leading to specific hazards like Haz06), user-based interfaces (e.g., an operator triggering the emergency stop), and vulnerability-based interfaces (e.g., the PLC being associated with vulnerabilities V4 and V5). These interfaces allow the formation of ILs, which connect related interfaces from different IEs, creating a traceable and interconnected system behavior and dependencies model.

Attributes play a fundamental role in supporting risk propagation and informed decision-making. Key attributes include version number, identity code, manufacturer details, failure rates, and installation or first-use dates for components. In this model, the user provides the date of first use or installation, assuming no replacements have occurred. For hazards, important attributes consist of identifiers, names, and the estimated probability of occurrence. These attributes ensure the model can support qualitative and quantitative analyses for comprehensive risk evaluation.

## 7.2 Introduction to BBN

A BBN is a probabilistic graphical model that facilitates reasoning under uncertainty by representing variables and their conditional dependencies. The two fundamental components of a BBN are the Directed Acyclic Graph (DAG) and CPT [72, 145, 144]. In the DAG structure, nodes represent random variables associated with a unique identifier and descriptive label.

Directed edges (arrows) denote causal or probabilistic dependencies between these variables. A node at the tail of a directed edge is called the parent, while the node at the head is the child. These directed relationships encapsulate the conditional dependencies between variables, which are quantified through CPTs [99, 71].

The overall joint probability distribution of the system can be factorized based on the structure of the DAG, as shown in Equation 7.1, where each variable is conditionally dependent on its parent nodes:

$$P(X_1, X_2, \dots, X_n) = \prod_{i=1}^n P(X_i \mid \text{parents}(X_i)) \quad (7.1)$$

### 7.2.1 Software tools for BBN development

Several software platforms and programming tools are available to support the development and quantification of BBNs. These tools enable users to build, analyze, and

simulate BBN models for various applications, including risk assessment, decision-making, and diagnostics [45]. Among the commonly used applications are:

- Hugin Expert <sup>1</sup>: A comprehensive decision support tool for modeling Bayesian Networks and Influence Diagrams. It offers commercial and academic versions and supports handling incomplete data, predicting future events, and performing diagnostic reasoning.
- Netica <sup>2</sup>: Designed for modeling Belief Networks and Influence Diagrams, Netica allows for defining relationships using direct probabilities, equations, or data-driven learning. Key features include support for dynamic assessment and no restriction on the number of nodes in the network.
- GeNIe <sup>3</sup>: A machine learning and probabilistic reasoning platform developed by the Decision Systems Laboratory at the University of Pittsburgh. GeNIe supports learning algorithms, influence diagrams, continuous probability distributions with equations, Dynamic Bayesian Networks (DBNs), and predictive and diagnostic inference. It is Windows-based and supports importing data from Microsoft Excel. The current implementation uses a static model, while dynamic modeling is planned as the next step.
- Analytica <sup>4</sup>: A platform for building Bayesian Networks and Influence Diagrams. Its primary limitation is the lack of support for backward (diagnostic) inference, as it only implements forward/predictive inference algorithms. It also lacks built-in support for computing updated probabilities based on new evidence.
- BayesiaLab <sup>5</sup>: An AI platform centered on Bayesian Networks that supports both forward and backward inference and dynamic Bayesian modeling. It is widely used in industrial, medical, and market research domains.
- Bayes Server <sup>6</sup>: A high-performance Bayesian Network software capable of risk modeling, sensitivity analysis, impact evaluation, and causal reasoning. It is particularly effective for reasoning under uncertainty in complex systems.

In addition to these applications, object-oriented programming environments such as Python offer several built-in libraries for developing and analyzing Bayesian Networks. Libraries like *pgmpy* and *bnlearn* enable users to model networks programmatically, integrate them into automated systems, and perform learning or inference using real-world datasets.

---

<sup>1</sup><https://www.hugin.com/>

<sup>2</sup><https://www.norsys.com/netica.html>

<sup>3</sup><https://www.bayesfusion.com/downloads/>

<sup>4</sup><https://analytica.com/products/free-analytica/>

<sup>5</sup><https://www.bayesia.com/>

<sup>6</sup><https://www.bayesserver.com/>

### 7.2.2 Relevance and Adoption of BBN in Risk Modeling

The application of BBNs for integrated risk assessment has gained momentum in research and industry. In [15], the potential for automating integrated risk assessments using BBNs is explored. The study compares various methods, including BTA, BDMP, and STPA-SafeSec, highlighting BBN's capability to model complex causal relationships.

According to [70], BBNs are particularly useful for modeling and analyzing uncertain systems. Their effectiveness has been demonstrated in several studies [146], with additional work like [74] addressing the challenge of data scarcity using fuzzy set theory.

Overall, the strength of BBNs lies in their ability to handle incomplete data, perform both predictive and diagnostic inference, support scenario analysis, and integrate diverse data sources, making them a powerful tool for combined safety and security risk assessment [45, 72].

### 7.2.3 Extensions of BBN

BBNs are well-suited for modeling complex risk scenarios with common uncertainty, incomplete data, and interdependent variables. Humans play key roles—operators, engineers, and maintenance personnel ensure safe and efficient operations [48, 22]. However, their actions can lead to errors with serious safety and financial consequences [2, 134]. Human error arises from organizational, supervisory, individual, and technical factors [48]. Security attacks can also trigger errors by manipulating technical elements like sensor data [124]. BBNs offers a probabilistic approach that allows these soft variables to be incorporated into a structured framework. One of our publications [22] explores the relationships between organizational, supervisory, individual, technical, and security factors and highlights its use for complete risk assessment.

BBNs can model multi-stage attack scenarios, where each attack step depends on the successful completion of previous ones. Nodes in the network can represent system components, vulnerabilities, attacker decisions, and detection events. The probabilistic nature of BBNs allows for incorporating factors like the likelihood of exploit success, attacker capability, or the presence of security controls. By updating the network in real-time as new evidence (e.g., intrusion detection alerts) becomes available, BBNs can support dynamic threat assessment and response. One of our publications [21] defines the relation between different vulnerabilities and enhances the overall risk assessment.

## 7.3 AML and BBN integration using Python

The data from the single source of truth (i.e. AML model) must be visualised and extracted for the Bayesian analysis. The assets and their extracted data must be further refined for BBN implementation. Over here, Python plays a significant role in serving as a common platform for integrating AML and BBN. Python is an ideal integration platform, offering a rich ecosystem of tools and libraries tailored for data extraction, processing, and probabilistic modeling.

We use Python to parse and process information from AML files, leveraging libraries such as ElementTree for efficient XML parsing. These tools allow structured data from the AML model to be loaded, extracted, and transformed to meet the input requirements of a BBN. This preprocessing step ensures that the information encoded in AML is appropriately aligned for probabilistic reasoning.

Python also supports robust libraries for building and managing Bayesian networks, such as PyMC3 and pgmpy. These libraries allow users to define the network structure, including nodes, dependencies, and CPTs, which can be derived from the data extracted from AML. Training the BBN can be done either through learning algorithms applied to available datasets or via manual specification of CPTs.

Once trained, the Bayesian network enables probabilistic inference—updating beliefs in light of new evidence. Python’s capabilities make carrying out these reasoning tasks straightforward and integrating the outcomes into the AML environment. For example, results from the BBN can update properties or attributes within the AML file or support decision-making processes based on inferred probabilities.

## 7.4 BBNs implementation

To use BBN for risk assessment, quantified probability values for the nodes of BBN need to be calculated. The quantification process uses formulae for the probability of failure for the component, the likelihood of occurrence for vulnerability, and the probability of human error. It is combined henceforth as Probability of Node ( $P(N)$ ).

### 7.4.1 Identified nodes for the model

Within the BBN, each node represents a key element of the system, characterized by its associated probability of occurrence. Each node contributes to the overall risk assessment and system reliability evaluation by capturing distinct failure modes and operational conditions. The following nodes have been identified for inclusion in the model:

- **Vulnerability Node:** This node represents the likelihood that a system vulnerability will be exploited. It is denoted as  $P(N)_{\text{vuln}}$  and is computed using Equation (7.5).
- **Component Node:** This node reflects the probability that a component will fail to operate under normal conditions. Represented as  $P(N)_{\text{component}}$ , it is calculated using Equation (7.2).
- **Human Node:** The probability associated with human error is derived from expert judgment and the Bayesian network model presented in [22]. For modeling purposes, expert estimates place this probability,  $P(N)_{\text{human}}$ , in the range of 2–10% annually.

- **Safety Hazard Node:** This node indicates the likelihood of a safety hazard occurring. If the system is malfunctioning, the probability of hazard occurrence,  $P(N)_{\text{hazard}}$ , is assumed to be 1.
- **Process Node:** Similar to the Safety Hazard Node, this node represents the probability that a given process will not be completed due to system malfunction. In such cases, the likelihood of process failure,  $P(N)_{\text{process}}$ , is also set to 1.

#### 7.4.2 Quantification of P(N)

##### P(N) of Component Node ( $P(N)_{\text{component}}$ )

Reliability refers to the ability of a system to operate safely and effectively over time. It is commonly quantified using the failure rate ( $\lambda$ ) or the MTBF [131]. In safety-critical systems, reliability parameters are often represented by Safety Integrity Levels (SIL) or Performance Levels (PL), as defined in IEC 61508 and ISO 13849-1, respectively. These standards typically express hourly failure rates, including SIL and PL classifications.

The probability of failure for a system component represents the likelihood that it will cease to function correctly due to factors such as aging, wear and tear, inherent faults, or random errors. This probability is mathematically expressed as:

$$P(N)_{\text{component}} = 1 - e^{-\lambda t}, \quad (7.2)$$

##### P(N) for Human Node ( $P(N)_{\text{HEP}}$ )

While often less predictable, human error probability (HEP) plays a significant role in system failures and safety outcomes. Responsibility for such errors involves individual operators, supervisory oversight, and broader organizational practices [48, 22]. In the context of ICS, human error can compromise safety and functionality, potentially leading to severe outcomes such as accidents, equipment damage, production downtime, financial losses, or even threats to human life [2].

In one of the prior studies [22], the probability of human error was modeled using a BBN, with expert input placing the annual likelihood of such mistakes ( $P(N)_{\text{HEP}}$ ) between 2% and 10%. This estimate accounts for the presence of human error in 60–70% of industrial accidents [2], although not all errors result in significant damage or loss. The selected probability range acknowledges the variability of human performance and the impact of mitigation strategies, providing a practical baseline for analysis while reflecting the complexity of real-world operational environments.

##### P(N) for Hazard Node ( $P(N)_{\text{haz}}$ )

In our design, a hazard node shows whether a hazard has happened based on the parent node. If a parent is an asset or another hazard, then the failure of the asset or the occurrence of the parent hazard can cause the child hazard to happen. If the Parent

node fails -> Hazard\_occurs is True, then  $P(N)_{\text{haz}}$  is 1. If the Parent node works -> Hazard\_occurs is False, then  $P(N)_{\text{haz}}$  is 0.

$$P(N)_{\text{haz}} = \begin{cases} 1, & \text{if parent node has failed} \\ 0, & \text{if parent node is operational} \end{cases} \quad (7.3)$$

#### **P(N) for Process Node ( $P(N)_{\text{process}}$ )**

An incomplete process may lead to an incomplete function of the overall system. Like a hazard node, a process node depends on its parent's state. Suppose a parent is an asset, a hazard, or another process. In that case, the asset's failure, the hazard's occurrence, or an incomplete parent process can also cause the child process to remain incomplete. If Parent node fails -> Process\_incomp is True, then  $P(N)_{\text{process}}$  is 1. If the Parent node works -> Process\_incomp is False, then  $P(N)_{\text{haz}}$  is 1.

$$P(N)_{\text{process}} = \begin{cases} 1, & \text{if parent node is operational} \\ 0, & \text{if parent node has failed} \end{cases} \quad (7.4)$$

#### **P(N) for the Vulnerability Node ( $P(N)_{\text{vuln}}$ )**

According to IEC 62443, a *vulnerability* is a weakness or flaw in a system's design, implementation, or operation that can be exploited to compromise its security. Vulnerabilities may exist at the component level or emerge from the interaction of multiple system elements.

To evaluate the severity and likelihood of vulnerabilities in system components, this study utilizes the (CVSS) <sup>7</sup>. Version 4.0 is the latest stable release; however, during the work, version 3.1 was the release available. It forms the basis for this analysis.

The CVSS framework is structured into three metric groups:

- *Base Metrics (BM)* – Describe the inherent characteristics of a vulnerability that remain constant over time and across different environments.
- *Temporal Metrics (TM)* – Capture attributes that change over time, such as the availability of exploits or remediation.
- *Environmental Metrics (EM)* – Reflects the operational context and how the vulnerability impacts a specific environment.

For determining the probability of vulnerability exploitation, denoted as  $P(N)_{\text{vuln}}$ , only the Base Metrics are considered to ensure a context-independent and consistent assessment.

<sup>7</sup><https://www.first.org/cvss/v3.1/specification-document>

Within the BM, four key elements—referred to here as the Exploitability Metrics significantly influence the likelihood of successful exploitation:

- *Attack Vector (AV)*: It represents the vulnerability exploitation possibility. The more remote an attacker can exploit the vulnerability, the higher the base score. It is categorized into four values:
  - *Network (N)* – The vulnerability is exploitable remotely over a network.
  - *Adjacent (A)* – The attacker must be on the same shared physical or logical network (e.g., Bluetooth, local subnet).
  - *Local (L)* – The attacker must have local access to the system (e.g., logged-in user).
  - *Physical (P)* – Physical interaction is required to exploit the vulnerability (e.g., inserting a USB).
- *Attack Complexity (AC)*: It describes complexity for a successful attack. Lower complexity results in a higher score. It includes:
  - *Low (L)* – The attack can be carried out without special conditions.
  - *High (H)* – The attack requires specific conditions that are not easily met.
- *Privileges Required (PR)*: It describes the necessity of privileges for successfully exploiting the vulnerability.
  - *None (N)* – No privileges are required.
  - *Low (L)* – The attacker requires basic user privileges.
  - *High (H)* – The attacker must have administrative or elevated privileges.
- *User Interaction (UI)*: It describes whether the exploitation of the vulnerability depends on the user's involvement.
  - *None (N)* – The vulnerability can be exploited without any user interaction.
  - *Required (R)* – Successful exploitation requires user participation (e.g., opening a malicious file).

Several studies have proposed methods to convert CVSS scores into probability values for use in probabilistic risk models such as BBN [18, 107, 143, 69]. Early work in [107] focused on CVSS v1.0, with later research extending conversion techniques to versions 2.0 and 3.0 [143, 69]. [18] introduced a refined approach for CVSS v3.1 by incorporating both base and temporal metrics. These efforts primarily focused on deriving  $P(N)_{\text{vuln}}$ .

In the present work, the Base Metrics are further divided into two distinct groups:

- **Exploitability Metrics:** used to calculate the probability of occurrence  $P(N)_{vuln}$  shown in 7.5.
- **Impact Metrics:** used to evaluate the Severity of Node ( $S(N)$ ),  $S(N)_{vuln}$  shown in 7.6.

Table 7.1 presents the quantitative values for each CVSS metric used in the computation of  $P(N)_{vuln}$ . These values are normalized to the  $[0, 1]$  range for compatibility with the probabilistic model.

Table 7.1: CVSS Metrics for Probability and Severity Calculations

Metric	Metric Value	Score
<b>Attack Vector (AV)</b>	Network	0.85
	Adjacent	0.62
	Local	0.55
	Physical	0.20
<b>Attack Complexity (AC)</b>	Low	0.77
	High	0.44
<b>Privileges Required (PR)</b>	None	0.85
	Low	0.62
	High	0.27
<b>User Interaction (UI)</b>	None	0.85
	Required	0.62

Decoupling probability and severity within the CVSS Base Metrics enhances the precision and interpretability of risk assessments. It enables more granular prioritization of vulnerabilities and supports informed decision-making in system protection and resource allocation. The  $P(N)_{vuln}$  is defined as

$$P(N)_{vuln} = AV \cdot AC \cdot PR \cdot UI, \quad (7.5)$$

### 7.4.3 Quantification of $S(N)$

In the context of a BBN, each component is modeled as a node whose failure can influence other network parts. The extent of this influence—termed severity—is determined by evaluating how many subsequent nodes are affected when a specific node fails. It enables a systematic assessment of cascading effects, where the failure of a single node propagates through its direct children and recursively through all connected descendant nodes. As such, the severity of a node reflects its topological importance and the breadth of its downstream influence within the network.

**S(N) of Vulnerability Node (S(N)<sub>vuln</sub>)**

The severity is assessed based on the vulnerability's potential impact on the CIA (Confidentiality, Integrity, Availability) triad, and is calculated using the following formula:

$$S(N)_{\text{vuln}} = 1 - [(1 - C) \cdot (1 - I) \cdot (1 - A)] \quad (7.6)$$

Here,  $C$ ,  $I$ , and  $A$  represent the respective impacts on confidentiality, integrity, and availability. The formulation ensures that a high impact in any one domain significantly increases the overall severity score.

**S(N) for all other Nodes**

For root nodes, those without any parent dependencies, the severity is considered at its highest. A failure at this level affects all downstream elements, making its influence network-wide. In contrast, non-root nodes, which depend on one or more parent nodes, generally exhibit a reduced severity since their influence is limited to their subtrees. The attenuation of impact is captured by modeling the severity score as the ratio of affected nodes to the total number of nodes in the network, accounting for diminishing influence as propagation moves deeper into the graph.

$$S(N)_{\text{component}} = S(N)_{\text{HEP}} = S(N)_{\text{haz}} = S(N)_{\text{process}} = \frac{\text{Number of affected nodes}}{\text{Total number of nodes}} \quad (7.7)$$

This hierarchical framework provides a straightforward method to quantify severity within the network. Nodes positioned higher in the structure possess a broader potential impact due to their control over larger network sections, while those situated deeper exert more localized effects. The model thus aligns the severity measure with the structural dependencies inherent in the BBN.

**7.5 Node Probabilities P(N) and Severity S(N)**

Table 7.2 presents the calculated metrics for various system nodes, including their probability of occurrence per hour P(N) and severity of impact S(N). Each row in the table corresponds to a unique node within the system, representing a component, sensor, vulnerability, or actor (such as a human operator).

The column P(N) indicates the likelihood of failure or event occurrence for that specific node within an hour, based on either failure rates, attack likelihoods, or empirical estimations. S(N) quantifies the severity of impact that the failure or event at the node would have on the overall system, considering the node's position and influence within the BBN structure.

Nodes such as *V3* and *V4* show relatively high risk values due to their high probability and significant severity combination. These nodes are likely to significantly influence the system's reliability and security, highlighting them as critical points of concern. The CVSS identified and represented in Table 5.1 are used to calculate the normalised  $P(N)_{\text{vuln}}$  (see 7.5) and  $S(N)_{\text{vuln}}$  (see 7.6) for the vulnerabilities.

On the other hand, nodes such as sensors (*e.g.*, *OSS*, *OSE*, *FLBS*) and human operators (*e.g.*, *OPR*, *MNC*) show significantly lower risk values. It indicates a lower probability of failure and/or a more localized impact on the system when such failures occur.

ID	Name	P(N) (per hour)	S(N)
V1	Brute force CBC	$3.0138 \times 10^{-1}$	0.43933
V2	Privilege Escalation CBC	$2.1431 \times 10^{-1}$	0.43933
V3	Denial of Service CBC	$6.8351 \times 10^{-1}$	0.48
V4	Denial of Service PLC	$9.4464 \times 10^{-1}$	0.48
V5	Code injection PLC	$2.0666 \times 10^{-1}$	0.9530
Comp	Compressor	$1.14 \times 10^{-5}$	0.4667
PLC	PLC functions as expected	$8.15 \times 10^{-6}$	0.6667
OSS	Optical Sensor at start	$1.60 \times 10^{-7}$	0.3
OSE	Optical Sensor at end	$1.60 \times 10^{-7}$	0.1
MTR	Motor for Conveyor belt	$1.76 \times 10^{-5}$	0.2667
LBSR	Light Barrier Sensor (red work-piece)	$1.60 \times 10^{-7}$	0.1667
LBSB	Light Barrier Sensor (black workpiece)	$1.60 \times 10^{-7}$	0.1667
FLBS	Forked Light Barrier Sensor	$1.60 \times 10^{-7}$	0.2333
SDRS	Sorting Deflector (red work-piece)	$3.85 \times 10^{-5}$	0.4
SDBS	Sorting Deflector (black work-piece)	$3.85 \times 10^{-5}$	0.4
CBC	Cobot Controller	$3.00 \times 10^{-6}$	0.4333
CBT	Cobot	$3.00 \times 10^{-6}$	0.4
OPR	Operator (Human)	$3.42 \times 10^{-6}$	0.3333
MNC	Administrator (Human)	$5.70 \times 10^{-6}$	0.5333
ES	Emergency Switch	$3.00 \times 10^{-6}$	0.0333

*Continued on next page*

ID	Name	P(N) (per hour)	S(N)
Safety01	Operator Injured	1	0.36667
Process01	Workpiece detected at start of conveyor	1	0.26667
Process02	Workpiece in downstream station	1	0.2333
Process03	Detection module determines workpiece color	1	0.2
Process04	Deflector for defined slide is advanced	1	0.16667
Process05	Workpiece ejected	1	0.1333
Process06	Deflector for defined slide is retracted	1	0.1
Process07	Workpiece detected at end of conveyor	1	0.06667
Process08	Workpiece picked up from end of conveyor	1	0.0333

Table 7.2: Calculated node probabilities  $P(N)$  and severities  $S(N)$ .

## 7.6 Individual Node Risk $R(N)$

The calculation of the risk for the node  $R(N)$  is performed based on the generated BBN. The risk is calculated based on the node's position in the network. Based on the CPT implementation, it would represent the node probability on the overall network. The  $R(N)$  would be different based on the implementation of the mitigation.

The first step in calculating the  $R(N)$  is identifying the risks. We have identified our use case's vulnerabilities, safety hazards, and processes. In this case, the risk will be related to exploiting vulnerabilities, occurring hazards, and the incomplete process. The probability value for the  $R(N)$  is determined for the system application using the BBN. The BBN inference method provides insights into different risk values for various scenarios.

Table 7.3 represents some of the identified safety and security risks. It also shows security risks that can compromise safety. For risk R001, the main risk nodes are  $R(N)_{MNC}$  and  $R(N)_{OPR}$ . The  $R(N)$  would be different if the risk is not satisfied (i.e., no mitigation) vs risk is mitigated. Also, for R002, Operator injury is dependent on sorting deflector functionality. In such a case, the  $R(N)_{OPR}$  is dependent on the risk probability of SDRS and SDBS (i.e.,  $R(N)_{OPR}$  if SDRS and SDBS malfunction or either malfunctions). In Chapter 9, further analysis on these outcomes is mentioned.

Risk ID	Risk name	Dependent Node	Risk Type
R001	Lack of Regular training for human	MNC, OPR	Safety
R002	Operator Injured due to Sorting deflector	SDRS, SDBS	Safety
R003	Operator Injured due to Cobot	CBT	Safety
R004	Operator Injured due to Compressor malfunction	Comp	Safety
R005	PLC unavailable due to DOS attack	V4	Security to Safety
R006	Attacker has higher privilege on Cobot controller	V2	Security to Safety
R007	Control of Cobot Controller compromised	V1	Security

Table 7.3: Identified risks for the use case

## 7.7 Relation between nodes

One of the fundamental principles of a (BBN) is the representation of dependencies between components using CPTs. These tables define the network's joint probability distribution, capturing how each node's state (a random variable) depends on its parent nodes [144]. In this work, nodes in the BBN model include physical components, vulnerabilities, safety hazards, and operational processes.

Nodes that do not have any parent nodes are assigned independent probabilities, denoted as  $P(N)_{node}$ . For dependent nodes, their behavior is governed by CPTs, which reflects the node's state's conditional probability given its parents' states. These conditional dependencies are structured according to specific relationship types between node categories. Table 7.4 summarizes the node-to-node relationships implemented in our model.

Fig. 7.3 provides a visual overview of the implemented relationships and CPT logic in the BBN. *Part A* of the figure illustrates the network structure using colored, ellipse-shaped nodes and directed edges labeled with relationship types. For instance, an arrow labeled "Vulnerability to Component" represents a dependency where a vulnerability node influences a component node.

Part B of Fig. 7.3 demonstrates an example CPT implementation for the relationship between a component and its vulnerabilities. The PLC node in this example has two associated vulnerabilities: V4 and V5. Additionally, a dependency of the PLC on OSS is shown. If the PLC fails, the conditional probability hinders the OSS function. If the PLC operates normally, the OSS follows its standalone probability  $P(N)_{component=OSS}$ . For instance:

Nr.	Parent Node	Child Node	Relation Description	Impact Description
1.	Vulnerability	Vulnerability	When both nodes represent vulnerabilities.	If the <i>parent</i> is exploited, the conditional probability of the <i>child</i> is $S(N)_{parent}$ . If the <i>parent</i> is not exploited, the <i>child</i> 's probability for not being exploited is $P(N)_{vuln}$ .
2.	Vulnerability	Component	A vulnerability leads to the compromise or malfunction of a component.	If the <i>parent</i> vulnerability is exploited, the conditional probability that the <i>child</i> component will be compromised is $S(N)_{parent}$ . If the <i>parent</i> is not exploited, the <i>child</i> probability follows its own risk $P(N)_{component}$ .
3.	Component	Component, Process, or Hazard	A component failure leads to functional failure of another component, process step, or safety issue.	If the <i>parent</i> fails, the conditional probability for <i>child</i> is $S(N)_{parent}$ . If not, the <i>child</i> behaves according to its own $P(N)_{component}$ .
4.	Human	Component or Process	Human errors propagate to system faults or process interruptions.	If the <i>parent</i> (human) commits an error, the <i>child</i> fails with conditional probability $S(N)_{parent}$ . Otherwise, the <i>child</i> 's conditional probability depends on <i>child</i> 's $P(N)$ .
5.	Process	Process	A sequential process dependency where the outcome of a process step depends on the completion of the previous step.	If the <i>parent</i> is incomplete, the <i>child</i> is incomplete with conditional probability of '1'. If the parent completes successfully, the child continues with probability $P(N)_{process}$ .

Table 7.4: Defined node-to-node relationships in the BBN model

- When both V4 and V5 are exploited (V4(0), V5(0)), the PLC is considered fully compromised. The CPT for PLC(0) with probability based on the  $S(N)_{V4}$  and  $S(N)_{V5}$ . The probability is  $\min((S(N)_{V4} + S(N)_{V5}), 1)$  to keep the probability value bound between 0 and 1.
- When only one vulnerability is exploited (V4(0) & V5(1) or V4(1) & V5(0)), the PLC(0) CPT is  $S(N)_{V4}$  or  $S(N)_{V5}$  depending on the vulnerability that is exploited.

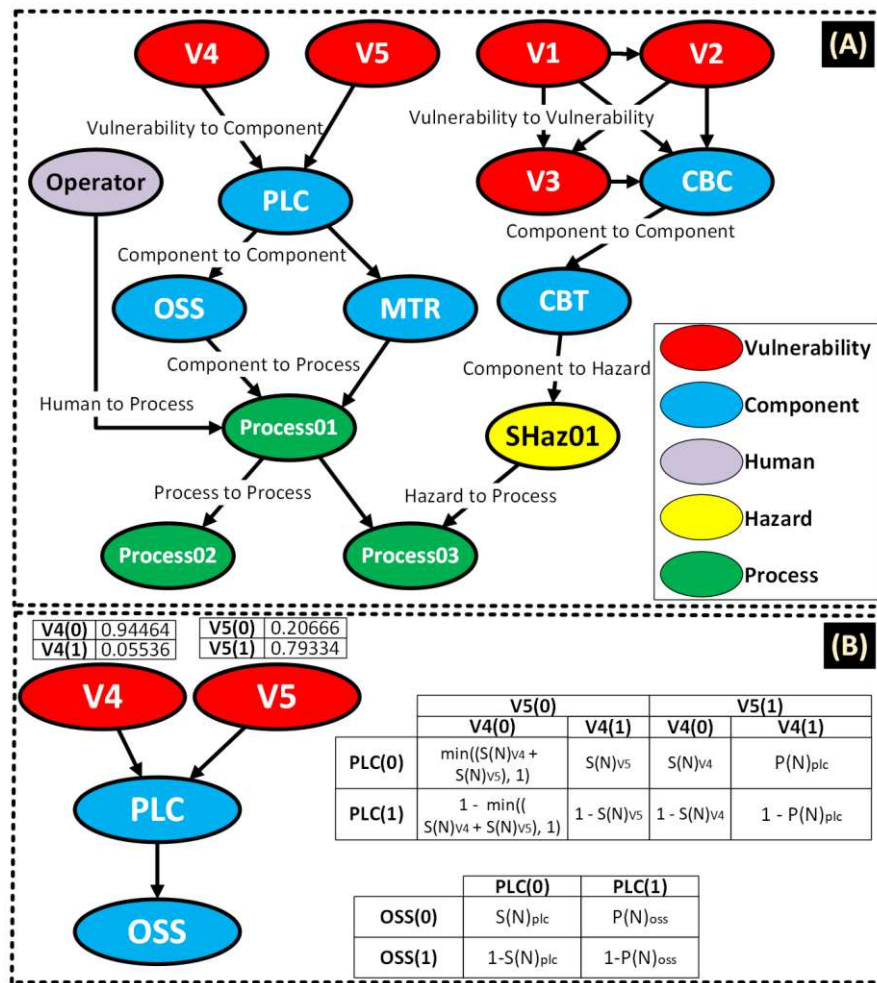


Figure 7.3: Representation of (A) Node-to-node relationships and (B) Example CPT implementation for the BBN.

- When neither is exploited ( $V4(1)$ ,  $V5(1)$ ), PLC failure is governed by its inherent failure rate  $P(N)_{component=PLC}$ .

## 7.8 CPT Assignment Algorithm for Child Node

The algorithms are designed to construct the CPT for a child node based on the type and number of its parent nodes. The CPT specifies the probabilities of the child node being in each possible state, conditioned on the states of its parent nodes. The key inputs to the algorithm are: the type of each parent node, the number of parents  $n$ , the severity  $S(N)$  associated with the child node, and the inherent probability of the child node  $P(N)$ . The output is a  $2 \times 2^n$  CPT, where each column corresponds to a unique combination of parent node states.

---

**Algorithm 7.1:** Conditional Probability Assignment for Child Node Based on Vulnerability or Component Parent
 

---

**Input** : Number of parents  $n$ , Severity of child node  $S(N)$ , Probability of child node  $P(N)$

**Output**: CPD for child node based on vulnerability or component parent

```

1 Initialize  $cpd[2][2^n]$  to zeros;
2 for  $i \leftarrow 0$  to  $2^n - 1$  do
3    $state \leftarrow$  Binary representation of  $i$  with  $n$  bits;
4   if  $n == 1$  then
5     if  $parent(0) == 0$  then
6        $cpd[0][i] \leftarrow S(N)$ ;
7     else
8        $cpd[0][i] \leftarrow P(N)$ ;
9     end
10     $cpd[1][i] \leftarrow 1 - cpd[0][i]$ ;
11  else
12    if  $state == all\ parents\ 0$  then
13       $cpd[0][i] \leftarrow \min(\sum S(N)_{parents}, 1)$ ;
14    else
15      if Only one parent is 0 then
16         $cpd[0][i] \leftarrow S(N)_{active\_parent}$ ;
17      else
18         $cpd[0][i] \leftarrow P(N)$ ;
19      end
20    end
21     $cpd[1][i] \leftarrow 1 - cpd[0][i]$ ;
22  end
23 end
24 return  $cpd$ ;

```

---

The algorithms initialize a CPD matrix of size  $2 \times 2^n$  with zeros. It then iterates over each possible combination of parent states, representing the binary form of integers from 0 to  $2^n - 1$ . Each bit in the binary representation corresponds to the state of a parent node, where 0 indicates that the parent is inactive (or failed) and 1 indicates that the parent is active (or functioning). Depending on the types of the parent nodes, the CPD values are calculated using different rules:

- **Vulnerability or Component Parent Nodes 7.1:**
  - If there is only one parent:
    - \* If the parent is in state 0, the probability that the child is in state 0 (safe or not failed) equals the severity  $S(N)$ .

**Algorithm 7.2:** Conditional Probability Assignment for Child Node Based on Process or Hazard Parent

---

**Input** : Number of parents  $n$   
**Output** : CPD for child node based on process or hazard parent

- 1 Initialize  $cpd[2][2^n]$  to zeros;
- 2 **for**  $i \leftarrow 0$  **to**  $2^n - 1$  **do**
- 3      $state \leftarrow$  Binary representation of  $i$  with  $n$  bits;
- 4     **if** *Both parents are 0* **then**
- 5          $cpd[0][i] \leftarrow 1$ ;
- 6     **else**
- 7         **if** *Only one parent is 1* **then**
- 8              $cpd[0][i] \leftarrow 1$ ;
- 9         **else**
- 10              $cpd[0][i] \leftarrow 0$ ;
- 11         **end**
- 12     **end**
- 13      $cpd[1][i] \leftarrow 1 - cpd[0][i]$ ;
- 14 **end**
- 15 **return**  $cpd$ ;

---

- \* If the parent is in state 1, the probability that the child is in state 0 is equal to its own probability  $P(N)$ .

- If there are multiple parents:

- \* When all parents are in state 0, the child's probability of being in state 0 is the minimum of the sum of the severities of the parents and 1.

- \* When exactly one parent is in state 0, the child's probability is determined by the severity of that particular parent.

- \* When all parents are in state 1, the child's probability is based on its inherent probability  $P(N)$ .

- **Process or Hazard Parent Nodes 7.2:**

- If all parents are of type Process or Hazard:

- \* If at least one parent is in state 0, the child is considered fully safe with probability 1.

- \* Only when all parents are in state 1 does the child node have a probability 0 of being safe, and thus a probability 1 of being in a failed or hazardous state.

- **Mixed Component and Process Parent Nodes 7.3:**

- If the child has a combination of component and process parent nodes:

---

**Algorithm 7.3:** Conditional Probability Assignment for Child Node Based on Component and Process Node
 

---

**Input** : Severity of child node  $S(N)$ , Probability of child node  $P(N)$

**Output** : CPD for child node based on component and process node

```

1 Initialize  $cpd[2][2^2]$  to zeros;
2 for  $i \leftarrow 0$  to 3 do
3    $state \leftarrow$  Binary representation of  $i$  with 2 bits;
4   if  $Component(0) \ \&\ \ Process(0)$  then
5     |  $cpd[0][i] \leftarrow 1$ ;
6   else
7     if  $Component(0) \ \&\ \ Process(1)$  then
8       |  $cpd[0][i] \leftarrow S(N)_{component}$ ;
9     else
10      if  $Component(1) \ \&\ \ Process(0)$  then
11        |  $cpd[0][i] \leftarrow 1$ ;
12      else
13        |  $cpd[0][i] \leftarrow P(N)$ ;
14      end
15    end
16  end
17   $cpd[1][i] \leftarrow 1 - cpd[0][i]$ ;
18 end
19 return  $cpd$ ;

```

---

- \* If both the component and process are in state 0, the child is safe with probability 1.
- \* If the component is 0 and the process is 1, the child's probability is determined by the severity  $S(N)$  of the component.
- \* If the component is 1 and the process is 0, the child is again safe with probability 1.
- \* If both the component and process are in state 1, the child's probability is its inherent probability  $P(N)$ .

After calculating the probability  $cpd[0][i]$  (child being in state 0), the complementary probability  $cpd[1][i]$  (child being in state 1) is set as  $1 - cpd[0][i]$  for each parent state combination. Finally, the fully populated CPT matrix is returned. This structured approach ensures that the child node's behavior accurately reflects the combined influence of its parent nodes, depending on their types and states.

The proposed algorithms offer several significant benefits. First, it provides a straightforward and automated method for generating CPT, eliminating the need for time-consuming manual expert elicitation for each node. Its highly adaptable design allows it to handle

various combinations of parent node types—such as vulnerabilities, components, processes, and hazards—making it applicable across a wide range of system architectures. Moreover, the algorithm ensures logical consistency by following a systematic set of rules for CPD generation, thereby maintaining the child nodes' coherent behavior relative to the parent nodes' influence. Finally, it exhibits strong scalability; programmatically evaluating all possible parent state combinations, the method remains effective even as the number of parent nodes increases, supporting applications for large and complex systems.

There are certain limitations to the above approach. It relies on relatively simplistic assumptions, such as additive or minimum-based relationships among parent nodes, which may fail to capture complex real-world interactions accurately. Additionally, the probability assignment is fixed based on predefined severities and does not adapt dynamically unless integrated with a separate learning mechanism. It limits the system's ability to respond to updated evidence. Furthermore, for systems characterized by nonlinear, synergistic, or antagonistic parent interactions, the expressiveness of the model may be insufficient to represent actual dependencies. However, as more empirical data or expert knowledge becomes available, these conditional probabilities can and should be refined.

The conditional or individual probability update can be integrated into the BBN graph. The complete graph can be updated with the up-to-date information supporting the uncertain nature of risks. Ultimately, this flexible framework supports the improvement of the iterative model. By capturing essential failure dynamics and then incorporating domain-specific insights, it enables qualitative and quantitative reasoning under uncertainty, making it a powerful tool for dependability analysis, fault diagnosis, and system resilience evaluation.

### 7.9 Program Flow and Representation

To extract and utilize data from an AML file, we employed Python's ElementTree library, which is included in the standard library and well-suited for parsing XML-based formats like AML. For the development and implementation of the BBN, we used the pgmpy library, a comprehensive Python toolkit for working with probabilistic graphical models. The overall workflow for the algorithm development is structured as follows:

1. *Import Necessary Libraries:* Use ElementTree for XML parsing and pgmpy to build and manage the BBN.
2. *Load the AML File:* Parse the AML file to gain access to its hierarchical structure. The "Sorting station for paper.aml" file is parsed using `ET.parse()` and the root element is obtained from the parsed XML.
3. *Extract relevant data:* Identify and extract useful data elements from the AML structure. IE elements within the XML using `root.findall()` are obtained. The ILS are extracted for node-to-node connections.

4. *Determine hierarchical relationships*: Analyze the extracted data to understand the parent-child relationships and system layout. IEs are the nodes in the BBN, while the ILs forms the directed edges between nodes. A list of combinations of parent nodes having multiple child nodes and multiple parent nodes for the common child node determines the connections.
5. *Define unique BBN nodes*: Translate key components from the hierarchy into distinct BBN nodes.
6. *Construct the **BBN** structure*: Establish dependencies and connections between nodes to form the directed acyclic graph (DAG).
7. *Assign conditional probabilities*: Specify the CPDs for each node based on domain knowledge or available data.
8. *Perform inference*: Use probabilistic reasoning to compute updated evidence-based beliefs.
9. *Integrate and analyze results*: Feed the inference outcomes back into the AML context for further decision-making or system adaptation.

An academic version of Genie software is used to represent the nodes and connections within the BBN. It is a limited, no-cost tool designed to implement BBN (see Fig. 7.4). It can also calculate the probabilities of the intended risk assessment by providing the assigned values to the nodes. It provides the risk propagation and also shows the dependencies between different nodes.

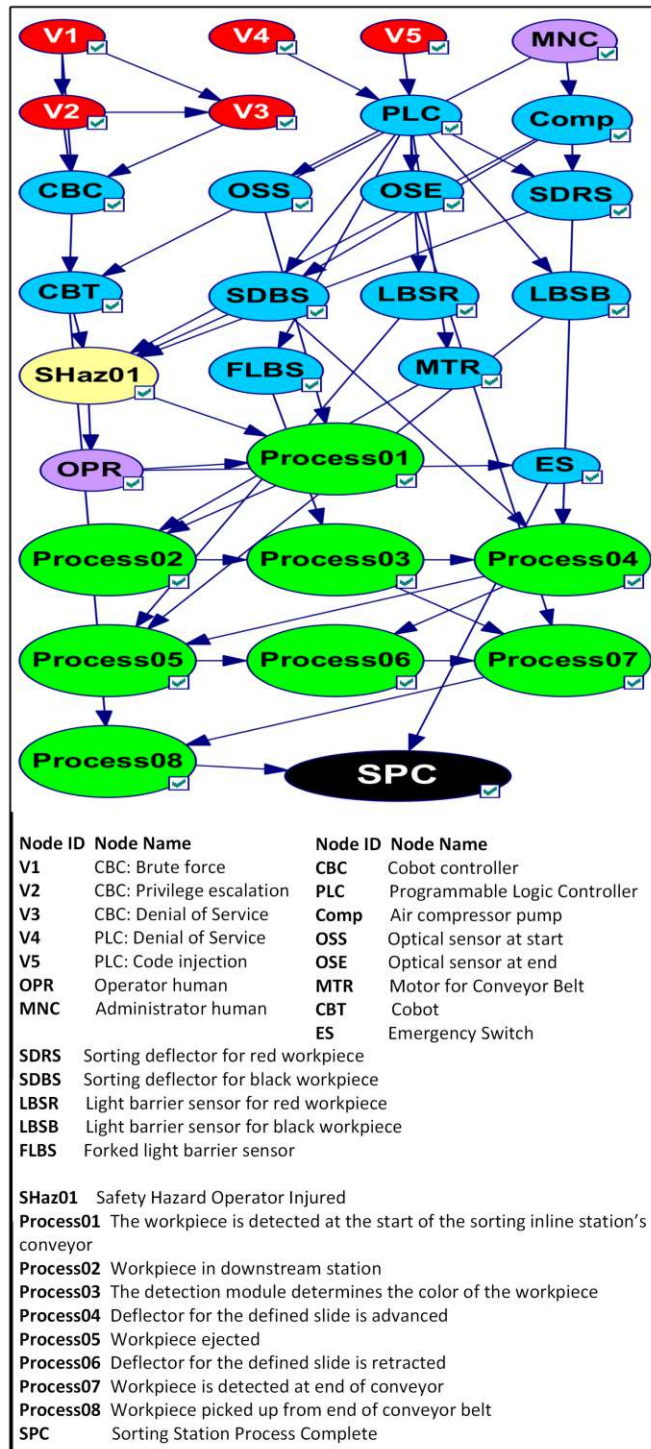


Figure 7.4: Representation of the BBN implementation.

# Risk Mitigation Strategy using Bayesian Belief Network

Risk treatment, a fundamental component of the broader risk management landscape, ensures complex systems' minimal or no risk operation, resilience, and sustainability. Understanding and managing risks is crucial for the smooth functioning of complex systems, especially in the world of ICSs. ICSs, which involves the control and automation of industrial processes, poses new challenges beyond safety concerns.

Risk treatment comprises four routes: mitigation, avoidance, transfer, and acceptance [3]. Depending on the measures usually in place in the organization, one or more of the four routes are considered. The research mainly focuses on the mitigation route based on the risk identification and assessment results. The choice of a route depends on factors like the economic aspect, technical capability, available tools, nature of the attack, and third-party contracts of the organization.

## 8.1 Sources for treatment measures

Risk mitigation strategies can be found and developed through a variety of sources. Industry standards and best practices often serve as foundational frameworks, offering recommended strategies and methodologies developed by industry associations, regulatory bodies, or international standards organizations. Regulatory guidelines play a significant role, providing specific requirements and expectations for risk management practices, particularly in industries subject to governmental oversight.

Consulting firms and risk management experts bring specialized knowledge, offering tailored advice based on industry expertise and experience. Case studies and industry reports provide valuable insights, allowing organizations to learn from the successes and failures of others when managing specific risks. Academic research, often published

by universities and research institutions, contributes to a deeper understanding of risk factors and effective mitigation strategies [46].

### 8.1.1 Standards as Sources for Risk Treatment in Safety and Security

Risk treatment in the safety and security domains requires a structured and systematic approach informed by internationally recognized standards. These standards provide guiding principles, methodologies, and requirements that assist engineers and decision-makers in identifying, assessing, and mitigating risks throughout the lifecycle of industrial systems. Each standard contributes uniquely to the overarching goal of system resilience.

IEC 61511 [10] focuses on functional safety in the safety industry and emphasizes protection mechanisms such as basic process control, prevention, mitigation, and emergency response. These mechanisms ensure that industrial processes remain within safe operating limits and that both plant personnel and surrounding communities are protected in case of hazardous events. Through structured layers of protection—from alarms and safety-installed systems to community emergency planning—IEC 61511 enables comprehensive risk control aligned with process safety goals.

ISO 12100 [59] outlines a foundational three-step method for machine safety. It starts with inherently safe design measures, safeguarding strategies, and comprehensive information. This hierarchy emphasizes the importance of designing hazards out of the system from the beginning and supports the implementation of technical and procedural measures to address any residual risk. This systematic approach enables designers to reduce hazards at the source while ensuring proper user awareness of remaining risks.

IEC 62443 [55] introduces cybersecurity principles into ICSs by segmenting systems into zones and conduits. This segmentation aids in risk containment and enhances system security by applying targeted controls based on the criticality and exposure of system components. The standard establishes a hierarchy of security requirements: foundational, system, and component, enabling a scalable approach to cyber risk mitigation. These are vital for safety-related systems, remote access, and wireless devices, which may be more exposed to threats.

NIST 800-30 R3 [121] further strengthens cybersecurity by promoting a defense-in-depth strategy incorporating technological, procedural, and organizational controls. It highlights the importance of network segmentation, continuous monitoring, and data integrity assurance, particularly within OT environments. The framework encourages separating corporate and industrial networks, using passive monitoring tools, and implementing robust incident detection and data protection mechanisms.

These standards collectively serve as a comprehensive foundation for addressing safety and security risks. They reflect ideal treatment approaches to guide initial system design and protection planning. However, it is essential to note that these relationships and frameworks often represent idealized assumptions or baseline best practices. As organizations gain deeper insights into their systems through real-time data, empirical

evidence, or operational feedback, these standards' conditional probabilities and risk treatment strategies can and should be refined. By continuously updating the risk treatment plans based on evolving knowledge, practical conditions, and new threats, practitioners can transition from theoretical modeling to more robust, context-sensitive decision-making, thereby improving safety, security, and system resilience in the long term.

### 8.1.2 MITRE ATT&CK and D3FEND Framework

MITRE ATT&CK Framework<sup>1</sup> looks at the problem from the adversary's perspective. It organizes the attacker's behavior in a series of tactics. Each tactic category mentions the techniques they employ to achieve the goal. It provides the user with information about adversaries related to:

- How did the information access the component?
- How did adversaries penetrate the network?
- How do they move laterally and gain privileged access?
- How do the adversaries evade the organization's security protocol?
- What is the attacker's goal, and what specific methods do they use?

MITRE D3FEND Framework<sup>2</sup> is a knowledge base, but more specifically, a knowledge graph, of cybersecurity countermeasure techniques. It is a catalog of defensive cybersecurity techniques and their relationships to offensive/adversary techniques. The primary goal of the initial D3FEND release is to help standardize the vocabulary used to describe defensive cybersecurity technology functionality. It will be released in 2021 and undergo various practical updates. However, some patents and techniques shown have many uses for the attack framework. One technique from defense can cause a significant improvement in the organization, as it can handle different attack frameworks.

### 8.1.3 Component Specific Mitigation

Vendors can play a crucial role in providing risk mitigation for their components by adopting proactive measures to enhance the security and resilience of their products. Making the component compliant with standards can provide a better option for the asset owner's safety and security experts. However, providing regular security updates is primarily applicable once the most diminutive owner has already used the component. It involves establishing a process for regularly releasing security updates and patches, promptly addressing and communicating about identified vulnerabilities, and providing customers with clear instructions on applying updates.

---

<sup>1</sup><https://attack.mitre.org/>

<sup>2</sup><https://d3fend.mitre.org/>

The asset owner needs to get updated if and when there is an updated feature for the used component. By exploring these various channels and resources, buyers can stay well-informed about security updates for the components they've purchased, enhancing the overall security posture of their systems. The common ways one can get this information are as follows:

- Vendor Website and Customer Support Information Documentation and Notifications
- Online Portals and User Accounts
- Security Bulletins and Announcements
- Automatic Update Features

### 8.1.4 Online Portals Providing Mitigation

Online portals such as the National Vulnerability Database (NVD), CVE Details, and others are crucial in enhancing organizational cybersecurity. These portals provide essential information about known vulnerabilities, enabling organizations to proactively assess and manage their digital landscape. By regularly monitoring these portals, security teams can get knowledge of the latest threats, understand the specific vulnerabilities affecting their systems, and take timely actions to mitigate potential risks. The detailed information provided, including severity metrics and mitigation strategies, empowers organizations to address vulnerabilities effectively, reducing the likelihood of exploitation.

Furthermore, these portals contribute to the overall cybersecurity posture by fostering collaboration and knowledge-sharing within the cybersecurity community. Organizations can leverage the collective intelligence on these platforms to gain insights into emerging threats, industry-specific risks, and best practices for safeguarding their systems. This shared awareness allows organizations to implement preventive measures, develop robust incident response plans, and fortify their defenses against evolving cybersecurity challenges. These online portals serve as valuable tools for organizations seeking to stay proactive, informed, and resilient in an ever-changing threat landscape.

Online portals that provide security-relevant data and best practices are as follows:

- **National Vulnerability Database (NVD)**<sup>3</sup>: is an online portal maintained by the National Institute of Standards and Technology (NIST) that serves as a central repository for CVE information. Each CVE entry in NVD includes detailed information about a vulnerability, including its description, affected software, severity level, and information. Significantly, NVD goes beyond listing vulnerabilities and provides mitigation strategies to help users reduce or eliminate the associated risks. It employs standardized severity metrics like the CVSS to assess and communicate

---

<sup>3</sup><https://nvd.nist.gov/>

the severity of vulnerabilities. With search capabilities, data feeds, and APIs, NVD enables security professionals to access timely updates, automate monitoring, and integrate vulnerability information into their security processes. In essence, NVD plays a crucial role in centralizing and disseminating vital information to empower cybersecurity professionals to make informed decisions and proactively manage systems.

- **CVE Database**<sup>4</sup>: The CVE database itself is a notable resource. It's a dictionary of publicly known cybersecurity vulnerabilities and exposures, and provides identifiers for each vulnerability. Various organizations use CVE identifiers to track and share vulnerability information.
- **SecurityFocus**<sup>5</sup>: SecurityFocus is a comprehensive cybersecurity portal that includes a vulnerability database. It provides information on vulnerabilities, exploits, and security news. Users can search for specific vulnerabilities and access related information.
- **Exploit Database (Exploit-DB)**<sup>6</sup>: Exploit-DB is a widely used online platform that provides information about exploits and vulnerabilities. It includes a database of exploits, shellcodes, and information papers, offering insights into potential security risks.
- **Vulnerability Lab**<sup>7</sup>: Vulnerability Lab is a platform that publishes information about security vulnerabilities, including details, advisories, and proof-of-concept. It covers a wide range of software and systems.
- **US-CERT (United States Computer Emergency Readiness Team)**<sup>8</sup>: US-CERT, operated by the Department of Homeland Security, offers security alerts, tips, and vulnerability information. It provides resources to enhance the nation's cybersecurity posture.
- **MITRE Corporation - CVE Details**<sup>9</sup>: CVE Details, maintained by MITRE Corporation, provides detailed information about specific CVEs. Users can search for vulnerabilities, view statistics, and obtain additional details such as vulnerability types and impact scores.
- **Packet Storm Security**<sup>10</sup>: Packet Storm Security is an online platform that offers a variety of cybersecurity resources, including a comprehensive vulnerability database. It covers vulnerabilities in software, hardware, and web applications.

<sup>4</sup><https://cve.mitre.org/>

<sup>5</sup><https://www.securityfocus.com/>

<sup>6</sup><https://www.exploit-db.com/>

<sup>7</sup><https://www.vulnerability-lab.com/>

<sup>8</sup><https://www.us-cert.gov/>

<sup>9</sup><https://www.cvedetails.com/>

<sup>10</sup><https://packetstormsecurity.com/>

- **SecuriTeam<sup>11</sup>**: SecuriTeam is a community-driven platform that provides information on vulnerabilities, exploits, and security news. It includes details on various security issues and vulnerabilities.
- **CISA (Cybersecurity and Infrastructure Security Agency) - Vulnerability Bulletins<sup>12</sup>**: CISA provides vulnerability bulletins that highlight significant vulnerabilities in various software and systems. These bulletins include details, risk assessments, and recommended mitigations.
- **GitHub Security Advisories<sup>13</sup>**: GitHub Security Advisories is a platform where researchers and maintainers can publish security advisories for their open-source projects. It lets users stay informed about security vulnerabilities in their software.

Online portals for safety services are valuable resources for organizations, offering comprehensive guidance and tools to enhance workplace safety. These platforms, such as OSHA, NIOSH, and others, provide informational safety regulations, standards, and best practices. Organizations can use these portals to stay informed about the latest safety guidelines tailored to their industry, helping them navigate and comply with regulatory requirements effectively.

Moreover, online safety portals offer practical resources for risk assessment, hazard identification, and developing robust safety management systems. They often provide downloadable tools, checklists, and educational materials that empower organizations to address potential safety concerns proactively. By utilizing the insights and recommendations offered by these portals, organizations can create a safer work environment, reduce the likelihood of workplace incidents, and prioritize the well-being of their employees. Additionally, the collaborative nature of these platforms fosters a community-driven approach to safety, allowing safety organizations to learn from industry best practices and shared experiences, ultimately contributing to a culture of continuous improvement in safety standards. Some of them are noted below:

- **Occupational Safety and Health Administration (OSHA)<sup>14</sup>**: The OSHA website provides information on occupational safety. It includes regulations, standards, guidelines, and information to help ensure a safe and healthy work environment. OSHA's Safety and Health Topics cover various industries and safety concerns.
- **National Institute for Occupational Safety and Health (NIOSH)<sup>15</sup>**: NIOSH, a part of the Centers for Disease Control and Prevention (CDC), offers resources and research on occupational safety. The NIOSH website provides guidelines, tools, and publications to address workplace hazards and promote worker well-being.

---

<sup>11</sup><https://securiteam.com/>

<sup>12</sup><https://us-cert.cisa.gov/>

<sup>13</sup><https://github.com/advisories>

<sup>14</sup><https://www.osha.gov/>

<sup>15</sup><https://www.cdc.gov/niosh/>

- **European Agency for Safety and Health at Work (EU-OSHA)<sup>16</sup>**: EU-OSHA Information and resources related to occupational safety and health in European workplaces. The information offers practical guidance, assessment tools, and best practice examples to enhance safety measures.
- **Health and Safety Executive (HSE) - UK<sup>17</sup>**: The HSE in the UK offers guidance on health and safety regulations, along with practical advice for implementing safety measures in various industries. Their website includes resources such as guidance documents, case studies, and risk assessment tools.
- **Centers for Disease Control and Prevention (CDC) - Workplace Safety and Health<sup>18</sup>**: The CDC's Workplace Safety and Health section provides information and resources to promote safety in the workplace. It covers various topics, including information preparedness, occupational health, and injury prevention.
- **National Safety Council (NSC)<sup>19</sup>**: The NSC is a nonprofit organization promoting safety in workplaces, communities, and homes. Their website offers resources, training materials, and tools to help organizations improve safety performance.
- **Institution of Occupational Safety and Health (IOSH)<sup>20</sup>**: IOSH, a professional body for health and safety practitioners, provides resources and guidance on safety practices. Their website includes publications, research, and tools to support organizations in creating safer working environments.
- **American National Standards Institute (ANSI)<sup>21</sup>**: ANSI develops and publishes safety standards for various industries. Their website provides access to a wide range of standards, including those related to safety practices. Standards often include recommended best practices for ensuring safety in safety contexts.

## 8.2 Risk Treatment Procedure

The process represented in Fig. 8.1 begins with the identification of the SUC using an existing AML model. The model provides the structural and operational context necessary for analyzing the system. Once the SUC is identified, the initial system risk is calculated using a BBN without including any mitigation nodes. The BBN represents the relationships between assets, hazards, and vulnerabilities in their unmitigated state. The data used for this initial risk calculation is derived from the attributes and relationships captured in the AML model, including asset properties, system topology, and known vulnerabilities.

<sup>16</sup><https://osha.europa.eu/>

<sup>17</sup><https://www.hse.gov.uk/>

<sup>18</sup><https://www.cdc.gov/niosh/topics/default.html>

<sup>19</sup><https://www.nsc.org/>

<sup>20</sup><https://www.iosh.com/>

<sup>21</sup><https://www.ansi.org/>

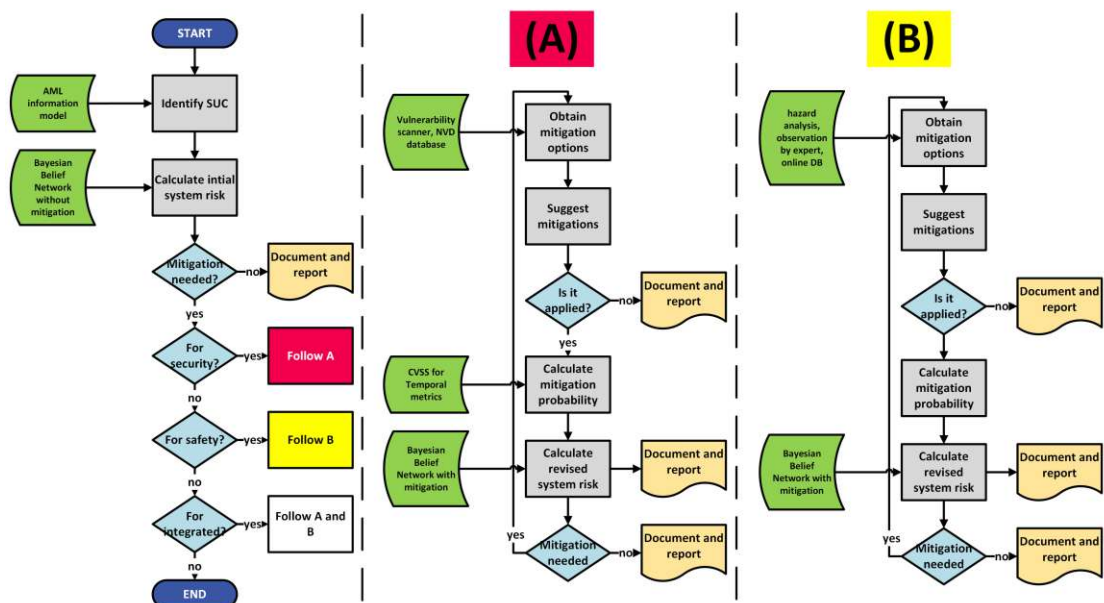


Figure 8.1: Risk treatment procedure for integrated approach

Following the initial risk calculation, a decision is made regarding the need for mitigation. If no mitigation is needed, the results are documented and reported. If mitigation is required, the next step is to determine the nature of the mitigation. If the issue is related to security, the process follows Path A. If it concerns safety, the process follows Path B. If the problem involves security and safety in an integrated manner, both paths are pursued. If none of these conditions apply, the process concludes.

For Path A, which addresses vulnerability nodes with a security focus, mitigation options are gathered from various sources such as vulnerability databases, vendor advisories, and standards. Suggested mitigations are then proposed based on the available options. A decision is made on whether the mitigation is applied. If not used, the situation is documented and reported. If the mitigation is applied, the mitigation probability, representing the effectiveness of the action, is calculated. The system risk is then recalculated using the BBN, this time incorporating the effects of the mitigation. Subsequently, a decision is made on whether further mitigation is necessary. If no further action is required, the results are documented and reported; otherwise, the process loops back to obtaining new mitigation options.

Similarly, Path B deals with hazard nodes from a safety perspective. Mitigation options are sourced from safety standards, industry best practices, and domain-specific guidance. After suggesting the appropriate mitigation, a decision is made regarding its implementation. If the mitigation is not applied, the outcome is documented and reported. The mitigation probability is calculated if applied, and the system risk is updated within the BBN. A final decision is made on whether additional mitigation is necessary. If not,

the process concludes with documentation and reporting; if yes, the process returns to sourcing additional mitigation options.

Throughout this process, mitigation decisions for security and safety are handled in an integrated manner, reflecting the interconnected nature of system risks. BBNs provides the flexibility to dynamically update system risks as integrated mitigations are applied, supporting a holistic, responsive, and iterative approach to risk management. Documentation and reporting at each decision point ensure traceability, transparency, and accountability across safety and security domains.

### 8.3 Identified mitigation for the risk

The mitigations for the vulnerabilities are obtained from the result of the vulnerability scanner. Also, if there are known CVEs for vulnerabilities, one can identify the mitigations from the NVD database. In Section 7.4.2, the Temporal Metric (TM) of the CVSS is mentioned. It defines the variable nature of vulnerability. One key Temporal metric is the Remediation Level (RL), which reflects the availability and maturity of fixes for a vulnerability.

Initially, most vulnerabilities are unpatched when disclosed. Over time, workarounds, temporary fixes, or official patches may emerge, progressively lowering the vulnerability risk. To add practicality for the mitigation effectiveness, probability values for official and temporary fixes are taken as 0.99 and 0.95, respectively. The workaround is assumed to be 80% effective (i.e., the probability value is 0.8).

Similarly, the mitigations for the safety hazard node depend on the impact the hazard causes. For practical purposes, we identify two mitigations. Train the user to reduce human error (SaM01) and properly maintain, repair, and replace the components and system (SaM02). There is no compromise in the implementation of safety mitigation. However, practically they are not 100% effective. Considering this, we assume the effectiveness to be more than 90%. The probability value for SaM01 is 0.95, and SaM02 is 0.9.

Table 8.1 presents a comprehensive mapping of identified system vulnerabilities to their corresponding mitigation strategies and effectiveness. Each vulnerability is uniquely identified (e.g., V1 for Brute Force, V2 for Privilege Escalation) and is paired with a mitigation identified by a code such as SeM01 or SeM02. For instance, the Brute Force vulnerability (V1) is addressed through mitigation SeM01, which involves limiting login attempts and is classified as an "Official Fix." Similarly, the vulnerability related to Denial of Service on PLCs (V4) is mitigated by upgrading to a higher software version, an "Official Fix." Other mitigation types include temporary fixes and workarounds, such as patching software for Privilege Escalation (V2) or restricting physical access to prevent Code Injection (V5).

Table 8.1 outlines various operational and cybersecurity risks and the corresponding mitigation measures applied. Each risk is uniquely labeled (e.g., R001 to R006) and

Table 8.1: vulnerability and its mitigations

Vul ID	Vul Name	Mitigation ID	Mitigation Description	Mitigation Effectiveness
V1	Brute Force	SeM01	Limit login attempts	Official Fix
V2	Privilege Escalation	SeM02	Patch software	Temporary fix
V3	DOS_CBC	SeM03	Implement traffic filtering and rate limiting	Workaround
V4	DOS_PLC	SeM04	Upgrade to higher version available	Official Fix
V5	Code Injection	SeM05	Restrict physical access to affected devices	Workaround

includes safety and security concerns. For example, risks such as operator injury due to sorting deflector, cobot, or compressor malfunction (R001–R003) are all mitigated using SaM02, indicating a shared safety mitigation strategy. Cybersecurity risks like PLC unavailability due to a DOS attack (R004) are mitigated by SeM04, which is consistent with the mitigation listed for vulnerability V4 in the previous table. This illustrates how a single mitigation can address a vulnerability and a broader system-level risk, highlighting the interconnection between vulnerability management and risk mitigation in critical systems.

Table 8.2: Risk and its mitigations

Risk ID	Risk name	Applied Mitigation
R001	Regular training for human	SaM01
R002	Operator Injured due to Sorting deflector	SaM02
R003	Operator Injured due to Cobot	SaM02
R004	Operator Injured due to Compressor malfunction	SaM02
R005	PLC unavailable due to DOS attack	SeM04
R006	Attacker has higher privilege on Cobot controller	SeM02
R007	Control of Cobot Controller compromised	SeM01

## 8.4 Integration of Mitigation node in BBN implementation

BBN can be used to model the relationships between different variables, such as the presence of a hazard or vulnerability, the effectiveness of mitigation strategies, and the outcomes of those strategies. In a Bayesian Network where vulnerabilities (e.g.,  $V4$  and  $V5$ ) affect a component (e.g., a PLC), integrating mitigations (e.g.,  $SeM04$  and  $SeM05$ ) can be approached in several ways:

### 8.4.1 Mitigations as Parents of Vulnerabilities

$$SeM04 \rightarrow V4 \rightarrow PLC, \quad SeM05 \rightarrow V5 \rightarrow PLC$$

This modeling assumes that the effectiveness of the mitigation ( $SeM04$  or  $SeM05$ ) directly influences the likelihood of successful exploitation of the respective vulnerabilities ( $V4$  and  $V5$ ). The CPT of  $V4$  and  $V5$  are adjusted according to the status of  $SeM04$  and  $SeM05$ , respectively. For example, if  $SeM04$  is active, the probability of  $V4$  being exploited is significantly reduced.

### 8.4.2 Mitigations as Parents of the Component

An alternative modeling approach is to connect mitigations directly to the component node ( $PLC$ ), alongside the vulnerabilities:

$$V4, V5, SeM04, SeM05 \rightarrow PLC$$

In this case, the mitigations are treated as factors that directly influence the  $PLC$ 's overall risk or operational status of the  $PLC$ , independent of individual vulnerabilities. This approach is suitable when mitigations impact the component as a whole (e.g., improving resilience broadly) rather than addressing specific vulnerabilities.

### 8.4.3 Introducing an Intermediate “Effective Vulnerability” Node

Another approach to integrating mitigation strategies into the Bayesian network is by introducing intermediate nodes that represent the *effective* vulnerability after considering mitigation. For each vulnerability, an intermediate node can be created that combines the effect of the vulnerability and its corresponding mitigation measure.

$$SeM04 \rightarrow V4\_eff \leftarrow V4$$

$$SeM05 \rightarrow V5\_eff \leftarrow V5$$

Here:

- $V4\_eff$  represents the *effective* state of vulnerability  $V4$  after considering the impact of mitigation  $SeM04$ .

- $V5_{eff}$  represents the *effective* state of vulnerability  $V5$  after considering the impact of mitigation  $SeM05$ .

After defining the intermediate nodes, the dependency of the PLC on vulnerabilities is modified:

$$V4_{eff}, V5_{eff} \rightarrow PLC$$

For the implementation in our BBN, we use the method shown in Section 8.4.2. Mitigation is applied to a vulnerable component. A similar strategy also implements the mitigation node for safety. It represents the practical applicability of the mitigation for the hazard or vulnerability.

## 8.5 Mitigation probabilistic relationship

Mitigation measures need to be applied effectively to both components and personnel. This means keeping software and firmware up to date for technical vulnerabilities, configuring systems securely, and performing regular maintenance checks. For safety hazards related to human errors, targeted training and implementing procedural changes are necessary. For instance, providing operators with detailed training on the operation of critical valves and implementing procedural checks can help prevent potentially hazardous mistakes.

There are safety and security mitigation nodes. The applicable mitigations for security are  $SeM01$  to  $SeM05$ , and for safety are  $SaM01$  and  $SaM02$ . The component with a vulnerability node will have a mitigation node for that vulnerability. The effectiveness of the mitigation will be represented in the CPT. For instance, if a component has one vulnerability node ( $V1$ ), one security mitigation node ( $SeM01$ ), and one safety mitigation node ( $SaM01$ ). The failure of the component or the impact of vulnerability will be reduced based on the mitigation effectiveness of the mitigation node.

$$V1, SeM01, SaM01 \rightarrow Component$$

The Table 8.3 outlines the conditional probabilities of a component failing based on three binary input variables:

- **SaM01** – Safety Mitigation Node Active (1: Active, 0: Inactive)
- **SeM01** – Security Mitigation Node Active (1: Active, 0: Inactive)
- **V1** – Vulnerability Exploited (0: Exploited, 1: Not Exploited)

Depending on the state of these inputs, different probability expressions are used to compute the likelihood of component failure ( $Component(0)$ ) and success ( $Component(1)$ ). The key quantities involved are:

SaM01	SeM01	V1	Component(0)	Description
0	0	0	$S(N)_{V1}$	Vulnerability exploited, no mitigation node, results in severity of vulnerability for component failure
0	0	1	$P(N)_{\text{component}}$	Vulnerability not exploited, no mitigations, results in Probability of failure for the component
0	1	0	$\max(\text{Mitigation Effectiveness (SeM01)} - P(N)_{V1}, 0)$	Vulnerability exploited, with security mitigation node active
0	1	1	$P(N)_{\text{component}}$	Vulnerability not exploited, Security mitigations, results in Probability of failure for the component
1	0	0	$S(N)_{V1}$	Vulnerability exploited, safety mitigation node, results in severity of vulnerability for component failure
1	0	1	Mitigation Effectiveness (SaM01)	Vulnerability not exploited, Safety mitigations, results in mitigation effectiveness
1	1	0	$\max(\text{Mitigation Effectiveness (SeM01)} - P(N)_{V1}, 0)$	Vulnerability exploited, with security and safety mitigations active
1	1	1	0	Vulnerability not exploited, Security and safety mitigations, results in 0 for failure of component

Table 8.3: Conditional Probability Table for Component Failure Based on Mitigations and Vulnerability

- $S(N)_{V1}$ : Severity of failure if vulnerability is exploited and no mitigation is in place.
- $P(N)_{\text{component}}$ : Probability of failure under normal (non-exploited) conditions.
- Mitigation Effectiveness (SeM01): Effectiveness of the security mitigation node.
- Mitigation Effectiveness (SaM01): Effectiveness of the safety mitigation node.

Each row in the table corresponds to a unique combination of inputs (SaM01, SeM01, V1), with the following logic applied:

1. **No mitigations (SaM01=0, SeM01=0):**

- If the vulnerability is exploited ( $V1=0$ ), failure probability is the severity  $S(N)_{V1}$ .
- If not exploited ( $V1=1$ ), failure probability is the normal probability  $P(N)_{\text{component}}$ .

2. **Security mitigation only (SaM01=0, SeM01=1):**

- If exploited ( $V1=0$ ), failure is calculated as  $\max(\text{Mitigation Effectiveness (SeM01)} - P(N)_{V1}, 0)$ .
- If not exploited ( $V1=1$ ), it behaves like the non-mitigated case.

3. **Safety mitigation only (SaM01=1, SeM01=0):**

- If exploited ( $V1=0$ ), failure remains  $S(N)_{V1}$ .
- If not exploited ( $V1=1$ ), failure probability is equal to the effectiveness of the safety mitigation.

4. **Both mitigations active (SaM01=1, SeM01=1):**

- If exploited ( $V1=0$ ), failure again follows  $\max(\text{Mitigation Effectiveness (SeM01)} - P(N)_{V1}, 0)$ .
- If not exploited ( $V1=1$ ), failure is guaranteed zero due to full mitigation coverage.

## 8.6 Representing BBN Graphs

To represent the nodes and connections within the BBN, we utilize Genie Academic software, a no-cost tool designed for implementing BBN (see Fig. 8.2 and 8.3). The visualization aids in comprehending the spread of risks as it represents the connection between components, their vulnerabilities, and their failure, leading to the hazards. Although the creation process is done manually, Python can also directly generate propagation visuals using libraries like *Matplotlib* and *NetworkX*. In our demonstration, we exhibited this process using Python. However, Genie Academic proves advantageous for enhanced representation and finer detail.

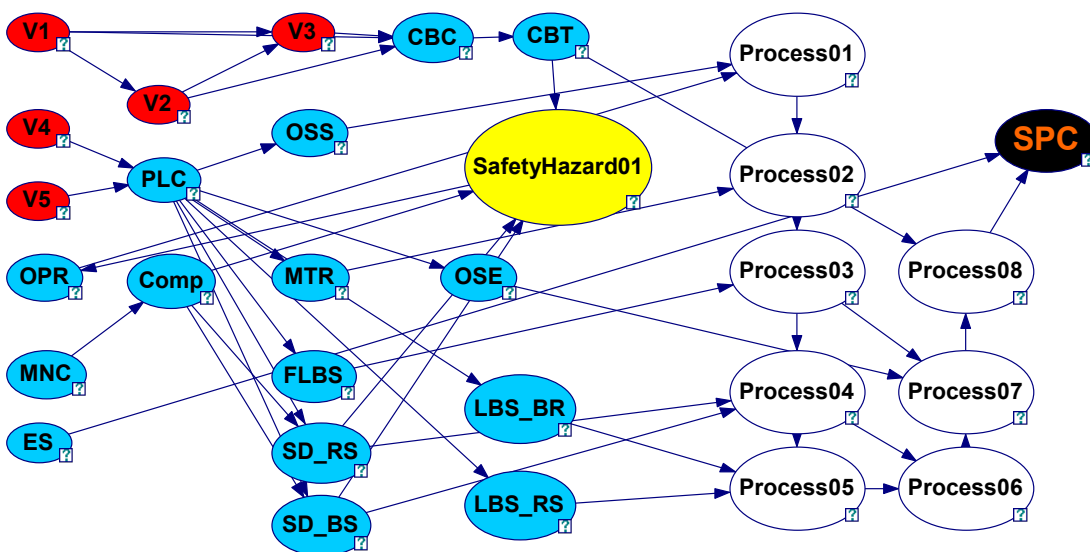


Figure 8.2: BBN risk propagation for use case without mitigation

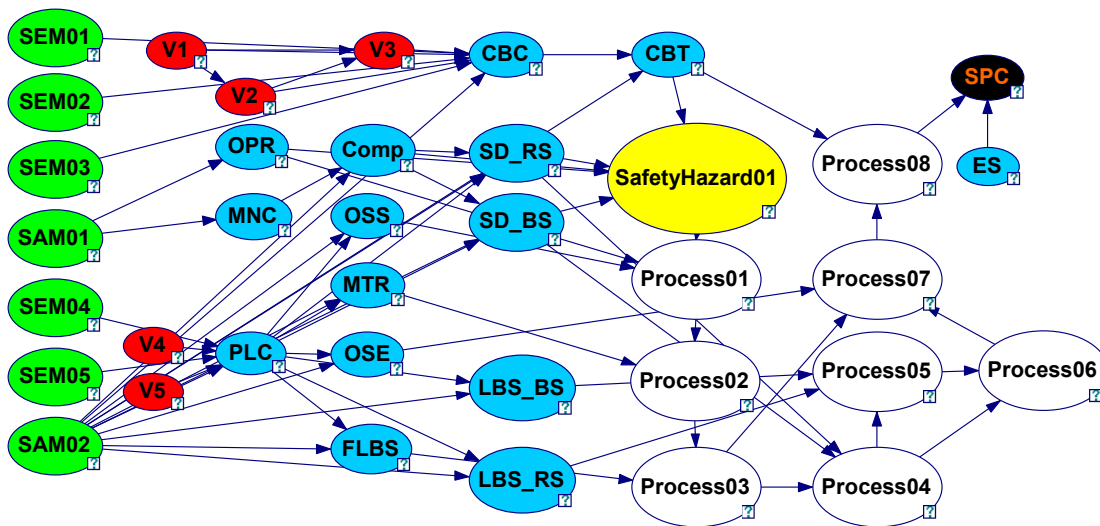


Figure 8.3: BBN risk propagation for use case with mitigation



# Results and Discussion

In this chapter, the outcomes of the integrated risk assessment and presented. The evaluation and interpretation of the results are discussed. It aims to demonstrate the effectiveness of the proposed methodology, interpret key findings, and highlight any limitations and areas for future improvement.

## 9.1 Risk assessment results

According to IEC 62443-3-2, the risk assessment strategy for the security should provide the assessment first without any mitigations in place. Then, perform the assessment with the mitigation implementation. The approach in the overall process is integrated safety and security. However, the strategy is similar to IEC 62443-3-2.

### 9.1.1 Risk Modeling and Input Configuration

The BBN provides a structured probabilistic framework to evaluate the risk of component failure in complex systems, incorporating both cyber and safety-related variables. Each node in the BBN represents a specific state variable, such as the activation of safety or security mitigations (SaM01, SeM01), or the exploitation state of a vulnerability (V1), while the edges define their causal influence on the failure behavior of the target component. The output of the BBN is a probabilistic quantification of component failure (Component(0)) or success (Component(1)) under various input configurations.

Additionally, the dynamic nature of the  $P(N)_{component}$  associated with system components is highlighted, emphasizing its fluctuation over time. The variability in  $P(N)_{component}$  values is considered within the risk calculation methodology. Although timely maintenance and repairs by mitigation SaM02 would provide measures, the component would, however, need replacements.

A user-centric approach to the display of risk results is presented, designed to cater to experts and asset owners by providing a comprehensive view of the system’s risk profile. User input, including the selection of a specific use case, the date of machine installation (with a default value provided), a reference date for risk assessment, and an estimation of successful attack probability, is incorporated to ensure that the risk assessment remains relevant and adaptable to evolving system conditions and available data.

**9.1.2 Probabilistic Risk Output and Comparative Analysis**

The risks for the individual nodes’ failure  $R(N)$  are calculated based on the BBN graph without mitigation and with mitigation. Ideally, according to ISO 31000, risk assessment should be performed periodically or with some changes in the system. The  $R(N)$  provides the overall understanding of the risk for the individual node.

In Table 9.1, the result of the obtained risks for the particular node is mentioned. The risks, with the dynamic nature of BBN graphs and  $P(N)$ , are calculated for the period of one year. The column provides the variation in the risk profile with and without mitigation. Overall system risk is also mentioned using the end node SPC.

Node ID	Node Name	R(N) (no mitigation)	R(N) (only safety mitigation)	R(N) (all mitigation)
Comp	Compressor	0.11668	0.009855	0.009855
PLC	PLC functions as expected	0.731292	0.7034	0.176411
OSS	Optical Sensor at start	0.487905	0.114717	0.027867
OSE	Optical Sensor at end	0.487905	0.114717	0.027867
MTR	Motor for Conveyor belt	0.526018	0.118528	0.03952
LBSR	Light Barrier Sensor (red workpiece)	0.487905	0.114717	0.027867
LBSB	Light Barrier Sensor (black workpiece)	0.487905	0.114717	0.027867
FLBS	Forked Light Barrier Sensor	0.487905	0.114717	0.027867
SDRS	Sorting Deflector (red workpiece)	0.598707	0.125354	0.053374
SDBS	Sorting Deflector (black workpiece)	0.598707	0.125354	0.053374
CBC	Cobot Controller	0.570028	0.501589	0.462384
CBT	Cobot	0.283694	0.079607	0.064181
OPR	Operator (Human)	0.60241	0.014755	0.008817

*Continued on next page*

Node ID	Node Name	R(N) (no mitigation)	R(N) (only safety mitigation)	R(N) (all mitigation)
MNC	Administrator (Human)	0.048836	0.002442	0.002442
ES	Emergency Switch	0.211144	0.005269	0.004346
Safety01	Operator Injured	0.590285	0.132011	0.070906
Process01	Workpiece detected at start of conveyor	0.643766	0.155266	0.076956
Process02	Workpiece in downstream station	0.684598	0.174806	0.083568
Process03	Detection module determines workpiece color	0.709596	0.190428	0.08739
Process04	Deflector for defined slide is advanced	0.779732	0.223887	0.101938
Process05	Workpiece ejected	0.798932	0.243171	0.106639
Process06	Deflector for defined slide is retracted	0.798932	0.243171	0.106639
Process07	Workpiece detected at end of conveyor	0.803416	0.24865	0.107974
Process08	Workpiece picked up from end of conveyor	0.815628	0.24865	0.11998
SPC	Sorting process completed	0.815824	0.260302	0.12002

Table 9.1: Calculated node risk probabilities  $R(N)$  for one year

Table 9.1 presents the computed risk values  $R(N)$  per hour for various nodes in the system under three conditions: without any mitigation, with safety mitigation, and with all mitigation strategies applied. These risk values are calculated by considering both the probability of a node failing and the severity of its impact on the overall system. The column titled  $R(N)$  (*no mitigation*) reflects the baseline risk when no risk-reducing measures are implemented. The second column,  $R(N)$  (*with safety mitigation*), accounts for improvements achieved solely through safety-oriented interventions such as emergency stop systems, operator awareness, and procedural controls. The third column,  $R(N)$  (*with all mitigation*), includes not only safety strategies but also technical and administrative mitigations like redundancy, fault-tolerant design, and automated diagnostics.

### 9.1.3 Interpretation of Risk Results

From the data, we observe a consistent and significant reduction in risk levels across almost all nodes as we move from the unmitigated state to the fully mitigated one. This

emphasizes the effectiveness of adopting a layered approach to risk reduction. Human-centric nodes such as the Operator (OPR) and Administrator (MNC) exhibit dramatic drops in risk, suggesting that interventions like training, ergonomic design, and human-machine interface improvements can greatly enhance system safety. For instance, the Operator's risk drops from 0.60241 without mitigation to 0.008817 with all mitigations applied.

The various process nodes, which represent stages in the material handling or sorting workflow, also show high initial risk values, typically above 0.64. These are progressively reduced through the introduction of safety measures and further minimized with full mitigation strategies, reaching values around or below 0.12. This demonstrates how critical process monitoring and control mechanisms are in improving system reliability.

Sensor and actuator nodes such as the Optical Sensors (OSS, OSE), Light Barrier Sensors (LBSR, LBSB, FLBS), and Sorting Deflectors (SDRS, SDBS) follow a consistent risk reduction pattern. Starting from approximately 0.48 in the unmitigated state, the risk values decrease to around 0.11 with safety mitigation and further down to 0.027867 with all mitigations in place. This reflects the importance of sensor redundancy and proper integration within control systems.

One particularly noteworthy pattern is observed with the PLC node. Here, the application of safety mitigations alone does not significantly reduce risk from 0.731292 to 0.7034. However, when safety and security mitigations are applied, including technical measures such as watchdog timers or dual-channel logic, the risk drops to 0.176411. It highlights the importance of robust system design for automation reliability.

The end node (SPC) provides the output of whether the *Sorting process is completed or not*. The probability value of SPC(0) and SPC(1) would provide the percentage of failures and success rates. To facilitate clearer interpretation and practical application of the quantitative outputs, the results could be translated into qualitative risk categories based on the success (i.e. state (1)) for the end node. These categories are defined as follows: Negligible (0–20%), Low (20–35%), Medium (35–50%), High (50–60%), and Critical (greater than 60%). The categorization provides a structured and accessible framework for assessing risk severity, thereby supporting effective prioritization of risk mitigation strategies.

Finally, the Sorting Process Completed (SPC) node, representing the successful conclusion of the operational sequence, begins with the highest risk at 0.815824, which is critical. With mitigation strategies, the risk decreases by more than 85%, ultimately reaching 0.12002. This reinforces the idea that ensuring process completion and reliability throughout all stages is central to system-level risk management.

### 9.1.4 Identified risk calculation

In the Python BBN library, the inference command enables users to compute posterior probabilities for selected variables by conditioning on observed evidence. These dependency calculations are governed by the CPTs (Conditional Probability Tables), which

define the probabilistic relationships between parent and child nodes in the Bayesian network. The inference process reveals how changes in one part of the system, represented by observed variable values, can influence the behavior of other components. Specifically, it captures the causal influence between variables (nodes) and the evidence (observed data) provided by the user. This is particularly useful in risk assessment, where understanding how certain failures propagate through the system is essential for designing effective mitigations.

Table 9.2: Risk and its mitigations

Risk ID	Evidence nodes	Node affected	R(N) (no mitigation)	R(N) (with mitigation)
R001	–	OPR, MNC	OPR(0) = 0.60241, MNC(0) = 0.048836	OPR(0) = 0.008817, MNC(0) = 0.002442
R002	SDRS: 0, SDBS: 0	OPR	OPR(0) = 0.8059	OPR(0) = 0.0769
R003	CBT: 0	OPR	OPR(0) = 0.4178	OPR(0) = 0.0466
R004	Comp: 0	OPR	OPR(0) = 0.4825	OPR(0) = 0.0647
R005	V4: 0	PLC	PLC(0) = 0.7621	PLC(0) = 0.1781
R006	V2: 0	CBC	CBC(0) = 0.8984	CBC(0) = 0.6874
R007	V1: 0	CBC	CBC(0) = 0.8633	CBC(0) = 0.6674

As illustrated in Table 9.2, a set of identified risks are quantified under different evidence scenarios. For instance, Risk R002: *Operator Injured due to Sorting Deflector*, quantifies the probability of operator failure under the condition that both sorting deflectors (SDRS and SDBS) malfunction, i.e., are in state 0. In this case, the calculated risk for the operator node ( $R(N)_{OPR}$ ) without mitigation reaches 0.8059, indicating a critical threat. However, once mitigations are applied—possibly including mechanical redesign, safety interlocks, or alert systems—the risk reduces significantly to 0.0769. This clearly demonstrates how Bayesian inference can quantify the impact of failure propagation and the effectiveness of mitigation strategies.

The risks, such as R003 and R004, show how failures in other nodes, like the cobot (CBT) and compressor (Comp), also influence the operator’s likelihood of injury and failure. These causal dependencies are not always intuitive, but the BBN framework makes them explicit, helping system designers uncover hidden risk pathways. For example, even a single point failure in a mechanical component (e.g., compressor) can increase operator risk significantly, unless mitigated properly.

The last three entries (R005 to R007) highlight cyber-related vulnerabilities (V1, V2, and V4) and their effect on programmable logic and controller nodes (PLC and CBC). These risks emphasize that both cyber and physical domains must be jointly considered in risk analysis. The elevated risk values in the absence of mitigation (e.g.,  $PLC(0) = 0.7621$  in

R005) and their substantial reduction post-mitigation ( $PLC(0) = 0.1781$ ) provide strong evidence of the utility of layered defenses.

The BBN framework not only calculates such risk values but also allows decision-makers to perform a sensitivity analysis, identifying which components most influence a given target node’s risk. This capability is invaluable for prioritizing risk reduction investments in complex systems. In sum, Bayesian inference serves as a powerful method for modeling, diagnosing, and mitigating risks in cyber-physical systems through a structured, probabilistic lens.

## 9.2 Evaluation based on standard ISO 31000

In Section 2.4, criteria for an overall risk assessment and management based on ISO 31000 are mentioned. The Table 9.3 provides a comprehensive evaluation of the integrated risk assessment method against the principles outlined in ISO 31000, the international standard for risk management. This evaluation highlights both the strengths and areas needing development in the context of implementing a BBN-based approach supported by AML within an industrial safety and security framework.

Table 9.3: Evaluation of the Risk Assessment Method Against ISO 31000 Principles

ID	ISO 31000 Principle	Satisfied	Explanation
(a)	Create value (human health and safety, product quality, financial value)	~ Partial	The method supports risk reduction and enhances human safety and reliability. However, direct linkage to product quality or financial impact is not explicitly modeled.
(b)	Be an integral part of organizational processes	~ Partial	The BBN-based method provides valuable insights, but an integration into operational or maintenance workflows and availability of a single source of truth needs to be demonstrated.

*Continued on next page*

ID	ISO 31000 Principle	Satisfied	Explanation
(c)	Be part of decision making (risk identification and treatment)	~ Partial	Supports decision-making by identifying risks and the effect of mitigations, but needs a stronger connection to actual decision-making protocols. Also, the model is suggestive and provides the analysis if certain nodes are implemented.
(d)	Explicitly addresses uncertainty (unknown situation)	✓ Yes	BBN are inherently probabilistic, capturing uncertainty through conditional probabilities and allowing inference under incomplete information.
(e)	Be systematic and structured	✓ Yes	BBN construction and inference follow a defined, structured methodology with nodes, edges, and conditional probability tables.
(f)	Be based on the best available information	✓ Yes	Utilizes data from the AML, integrating, to our knowledge, up-to-date technical and operational input.
(g)	Be tailored	✓ Yes	The model is built for specific system configurations and use cases, reflecting system-specific risks and components. However, the model can be generalised for many industrial applications.
(h)	Take into account human factors (internal and external risk actors)	✓ Yes	Human roles such as operators and administrators are included in the network as nodes, and their risk is modeled directly.

*Continued on next page*

ID	ISO 31000 Principle	Satisfied	Explanation
(i)	Be transparent and inclusive	~ Partial	The methodology is based on obtaining the probability of failures for the components, which is not transparent or sometimes not known. BBN can operate with approximates and provide updates to the overall system when new information is available.
(j)	Be dynamic, iterative, and responsive to change	✓ Yes	The BBN allows updates based on new evidence, making it adaptable to changing system conditions.
(k)	Be capable of continual improvement, update, and enhancement	✓ Yes	The model can be extended with new data, evidence, and updated CPTs, supporting continuous improvement.

Firstly, the method clearly satisfies several foundational principles. It explicitly addresses uncertainty, which is one of the core strengths of BBNs, as these models operate inherently on probabilistic logic. This allows them to perform inference even when information is incomplete or uncertain. Similarly, the method is both systematic and structured, given that BBNs are formally constructed using nodes, edges, and CPTs. The method also demonstrates tailoring to specific industrial configurations and systems, while remaining general enough to be adapted to various scenarios. Furthermore, human factors are directly modeled through the inclusion of roles such as operators or administrators, thus enabling a realistic risk representation that includes internal and external human actors.

The model is also dynamic and responsive to change, given its ability to update node probabilities when new evidence becomes available. This responsiveness extends to its support for continual improvement, as CPTs and system structure can be revised over time to reflect changes in the system or its environment. These characteristics show that the method is well-positioned for use in environments where change is constant and adaptability is essential.

### 9.3 Evaluation based on automation capability

Automation plays a critical role in today’s world. If implemented correctly, it provides greater assistance to humans in decision-making. It also helps in completing tasks in a timely and efficient manner. However, the automation can be based on various levels

[117, 44, 94]. The automation pyramid can also be a layer of automation increased based on the level, the component is placed on [115].

The level of automation in our case would come from the implementation of the integrated risk assessment and mitigation process. Each process is outlined, and the level of automation is mentioned. Also, if the process is manual, then the automation implementation is suggested as a future aspect.

### 9.3.1 Automation Levels Definition

Automation levels describe the degree to which a task is performed independently by a system without human intervention [117, 94, 44]. In this context, we apply a 0–5 scale presented in the [44]:

- Level 0 – Fully manual: All decisions and actions are human-driven.
- Level 1 – Assisted: The system provides guidance (e.g., templates or suggestions), but humans perform actions.
- Level 2 – Semi-automated: The system automates subtasks but relies on humans for supervision or control.
- Level 3 – Conditional automation: The system can execute specific processes based on conditions or user initiation.
- Level 4 – High automation: The system performs tasks autonomously within a defined scope and maintains internal consistency.
- Level 5 – Full autonomy: The system operates independently in dynamic contexts with learning and adaptation.

### 9.3.2 Process Description with Automation Levels

The integrated safety and security risk assessment process begins with the information collection phase, which currently operates at automation level 0. In this step, relevant data sources are manually identified based on expert knowledge, and information is gathered through documentation reviews, stakeholder consultations, and manual system analysis. These actions require significant human involvement with no system assistance.

In the information modelling phase, the gathered data is manually structured into formal representations using standards such as AML, AAS, and Ontologies. The process is guided by pre-designed templates, which assist users in ensuring uniformity and correctness, representing automation level 1. Though helpful, this assistance still requires humans to input and interpret all data.

Once modelling is complete, the information extraction phase demonstrates a marked increase in automation. A Python-based integration tool extracts data from AML files

automatically, identifying relevant components and relationships. It maps these to BBN nodes and directed edges without manual intervention—this step reaches automation level 4.

Additionally, the tool continuously monitors the AML file for updates (such as probability changes or structural modifications), and any changes are automatically reflected in the BBN. This real-time synchronization ensures consistency and adaptiveness within a predefined scope, another indicator of level 4 automation.

In the risk modeling and information usage phase, the system continues to perform at a high level of automation. Algorithms generate CPT based on the extracted data (level 3 automation), and inference engines compute both nodal and system-level risks (level 4). The system delivers outputs automatically, but users may still be involved in validating or interpreting results, and in some cases, responding to the output with additional inputs (level 2).

### 9.3.3 Overall Automation Level Estimation

Considering the entire pipeline—from manual data sourcing to automated risk assessment—the system achieves an overall automation level of 3 (conditional automation). This reflects a balance between fully automated components (such as AML parsing and BBN updating) and those that remain manual or semi-automated (like data gathering and model population).

The low automation at the front-end (collection and modelling) limits the system's autonomy across diverse environments. The high automation in the core integration and inference processes significantly enhances consistency, scalability, and responsiveness to changes.

The system demonstrates the ability to act independently within a defined context (AML-BBN linkage), but still depends on humans for initiating processes and handling broader context variability.

To progress toward Level 4 (high automation) overall, automation could be introduced in earlier stages, such as intelligent data source discovery, automated form filling using NLP or OCR, or automated selection of modeling techniques based on system type.

## 9.4 Advantages of the Integrated Risk Assessment Method

Beyond aligning with ISO 31000's foundational principles, the proposed risk assessment framework exhibits several advanced attributes. It is probabilistic, quantitative, and model-driven, with potential for partial or full automation [15]. These characteristics offer the following key advantages:

- **Holistic Risk Coverage:** The approach enables comprehensive treatment of diverse risk categories—including safety, security, human, cyber, and process-related

threats—in a unified framework. As shown in Fig. 8.3, the model incorporates a wide range of node types (hazards, assets, process states, vulnerabilities, and mitigation) through modular representation, thereby ensuring system-wide coverage.

- **Causal Dependency Analysis:** Leveraging the core properties of BBN, the method facilitates cause-and-effect reasoning through conditional dependencies [42]. This is evident in Fig. 7.4, where sensor failure probabilities are influenced by both upstream controller reliability and the component’s own characteristics, enabling transparent tracing of fault propagation.
- **Efficiency through Automation:** Integration with AML enables structured, automated extraction of relevant system data [147], supporting faster setup and more consistent information flows. Although AML requires manual updating, the BBN component can dynamically adapt to new evidence or conditions [16], making the inference process significantly more responsive compared to static methods [135].
- **Accurate Risk Estimation:** The synergy between AML and BBN allows for precise modeling of uncertainties and interdependencies, which is critical in domains where high uncertainty can lead to high loss potential [42, 65]. This leads to more reliable risk quantification.
- **Support for Informed Decision-Making:** Through graphical modeling and probabilistic reasoning, the system delivers interpretable outputs, allowing for both predictive (forward) and diagnostic (reverse) analysis [65, 25]. This dual capability aids operational and strategic decisions by illustrating likely causes and consequences under varying scenarios.
- **Modular and Scalable Architecture:** The modular nature of BBN facilitates easy integration of additional nodes, assets, or subsystems. Modifications in the underlying AML structure (e.g., the addition of new equipment) are seamlessly reflected in the BBN model, ensuring scalability and iterative model development [65]. The ability to merge BBNs with overlapping nodes supports collaborative or phased modeling efforts.

#### 9.4.1 Limitations of the Proposed Approach

Despite its strengths, the combined AML-BBN risk assessment method has certain constraints that warrant consideration:

- **AML Standardization and Static Nature:** Although AML offers a standardized representation of system architecture, the modeling of safety and security data is not universally standardized. User-specific extensions and interpretations can affect model quality. Moreover, AML serves as a largely static information source unless manually maintained.

- **User Dependency and Expertise Requirements:** Effective use of the methodology depends heavily on user expertise in both system engineering and risk modeling. Constructing accurate AML models and valid BBN networks requires substantial domain knowledge and familiarity with probabilistic reasoning.
- **Design Complexity and Maintenance Effort:** Developing intricate information hierarchies in AML and setting up robust Bayesian structures can be complex and time-consuming. This challenge is particularly prominent in large-scale or legacy systems where component relationships and dependencies are not readily documented.
- **Interpretability and Cognitive Load:** Interpreting the output of a complex BBN can be cognitively demanding, especially when multiple interconnected risks and mitigation scenarios are evaluated. This may limit the accessibility of the results to stakeholders unfamiliar with probabilistic inference.
- **Data Integrity and Organizational Alignment:** Integrated models assume reliable and complete data input. In practice, ensuring data integrity across different organizational units can be challenging. Additionally, harmonizing risk perspectives across safety, security, and operations requires a culture of cross-functional collaboration.
- **Potential Bias and Assumption Dependency:** As with any model-based approach, there is a risk of embedded assumptions or biases in node definitions, CPTs, or model boundaries. A transparent and iterative validation process is necessary to mitigate such concerns and ensure credibility.

Overall, while the proposed integrated method aligns well with international risk management standards and demonstrates technical robustness, its full potential is realized only when applied within an organization that supports structured modeling, interdisciplinary cooperation, and continuous learning [45].

## Conclusion and Future Work

The research presents the necessity and development of an integrated safety and security risk assessment approach tailored for ICS. It proposes a structured and unified methodology that holistically evaluates both safety and security risks, thereby illuminating the interdependencies and potential cascading effects within ICS components. A practical implementation is demonstrated using a sorting station setup comprising controllers, a collaborative robot (cobot), and various field devices.

Traditional assessment methods often fail to account for the dynamic dependencies and uncertainties inherent in complex ICS environments. Our approach leverages a BBN-based methodology, enabling the incorporation of probabilistic reasoning to represent both uncertainty and dependency. It is crucial for modeling how one failure or vulnerability can influence others across the system.

In addressing the practical needs of stakeholders, the research integrates advanced information modeling techniques—specifically, AML, AAS, and semantic ontologies. These models enable the systematic collection, structuring, and retrieval of heterogeneous safety- and security-relevant data. A key contribution of this research is the modeling of diverse information sources into a unified "single source of truth," facilitating more efficient, automated, and accurate risk assessment.

Python serves as the unifying platform for integrating the information models with the risk assessment methodology, supporting the generation of meaningful outputs such as node probabilities  $R(N)$ , which helps identify critical risk-contributing components in the system. It demonstrates the feasibility and benefits of integrated risk identification, analysis, and treatment in ICS environments.

The research answers several research questions mentioned in Chapter 1:

1. *What are the data input sources for safety and security-relevant data applicable in ICS?*

-> The identified data input sources for the research are interviews with a defined questionnaire, system architecture and asset-related information, process definition, penetration or vulnerability scanning documents, and implemented automated tools like IDS or SIEM. In the current research, we have implemented stakeholder analysis and vulnerability scanning on the use case.

2. *What are the safety and security risk indicators based on the collected data?*

-> The methodology of risk assessment is based on the probability of occurrence  $P(N)$ , severity of occurrence  $S(N)$ , and risk of occurrence  $R(N)$ . The information collected provides the necessary  $P(N)$ , while the  $S(N)$  is inferred from the architecture.  $R(N)$  are calculated based on the BBN graph. Other than that, incident history, system redundancy levels, human error, and implemented mitigation also provide necessary risk indications.

3. *Which information models are most suitable for supporting the automation of risk assessment and management in ICS?*

-> In the research, we have used and compared AML, AAS, and Semantic Ontology in Chapter 6. The most effective models are AML for describing plant engineering structures, focusing on engineering data, AAS for representing digital twins of physical assets and focusing on operational and engineering data, and Ontologies (OWL-based models) for capturing semantic relationships and reasoning over system behavior and threats. However, for the implementation of risk assessment, based on static ICS, AML is implemented.

4. *Which methods are suitable for an integrated safety and security risk analysis and assessment?*

-> In the research, we compare various risk assessment methods in Table 2.3. These include safety, security, and integrated risk assessment methods. After careful consideration, we choose BBN as it offers probabilistic dependencies. It also handles uncertainties and can integrate safety and security analysis. It can handle complexity and has the possibility for automation. It provides the capability to update based on new information.

5. *What methodology can be used to identify and implement risk treatment strategies?*

-> The methodology to implement risk treatment for the identified hazards and vulnerabilities is represented in Fig. 8.1. The proposed methodology includes BBN-based scenario analysis, and Chapter 8 uncovers the implementation strategy.

The evaluation framework aligns with the principles outlined in ISO 31000 for risk management. Out of the total requirements, our method completely satisfies two-thirds of the requirements, while the remaining requirements are partially satisfied. The partial satisfaction is based on the application usage and future requirements of the implementation.

The level of automation in the implementation is evaluated against automation maturity in [44]. Information collection remains fully manual (Level 0), while information modeling

---

is assisted via structured templates (Level 1). Significant automation begins during information extraction and BBN generation, which achieves Level 4 due to autonomous data parsing, node creation, and model updates. Conditional probability generation (Level 3) and risk inference (Level 4) further demonstrate automation maturity.

The research provides a robust, adaptable methodology for integrated safety and security risk management in ICS. It emphasizes data-driven modeling, cross-disciplinary integration, and automation support, enabling more resilient and informed decision-making processes.

Currently, the risk assessment implementation is semi-automated. The process of data collection and organization remains fully manual. Integrating automated data collection and machine learning for the identification of key risk indicators can be implemented as part of enhancements. The machine learning capabilities can also be used for CPT definition as part of Level 5 automation.

Risk in monetary terms is better understood by management. As a future aspect, the implementation of semantic reasoning to incorporate business process risks, supply chain threats, and organizational factors in addition to technical safety and security.

Current implementation depends mainly on the static nature of information. However, BBN can handle varying data. Future implementations may couple the AML-based model with digital twin architectures of AAS to simulate failure, attack, and mitigation scenarios under varying conditions. It currently provides partial predictive maintenance, however, complete predictive maintenance with conditional varying CPTs and real-time decision support based on live system behavior can be implemented.

Although BBNs are capable of representing complex and even non-linear dependencies, the present implementation makes use of relatively simple assumptions to demonstrate feasibility. Future work can be focused on extending the approach to leverage the full potential of BBNs (and complementary techniques) so that it can also address more intricate, non-linear relations in dynamic environments.



# List of Abbreviations

- AAS** Asset Administration Shell
- ALARP** As Low As Reasonably Practicable
- AML** Automation Markup Language
- API** Application Programming Interface
- ATA** Attack Tree Analysis
- BAD** Behavioral Anomaly Detection
- BBN** Bayesian Belief Network
- BDMP** Boolean logic-driven Markov Processes
- BTA** Bow Tie Analysis
- CAEX** Computer Aided Engineering Exchange
- CI** Critical Infrastructure
- CPS** Cyber Physical System
- CPT** Conditional Probability Table
- CVSS** Common Vulnerability Scoring System
- CVE** Common Vulnerabilities and Exposures
- DB** Database
- DOS** Denial of Service
- EI** External Interface
- ERP** Enterprise Resource Planning
- FMEA** Failure Modes and Effects Analysis

**FTA** Fault Tree Analysis

**GTST-MLD** Goal Tree Success Tree - Master Logic Diagram

**HAZOP** Hazard and Operability Methodology

**HMI** Human Machine Interface

**HRA** Human Reliability Analysis

**I4.0** Industry 4.0

**IACS** Industrial Automation and Control System

**ICS** Industrial Control System

**ICL** Interface Class Library

**IE** Internal Element

**IH** Instance Hierarchy

**IIoT** Industrial Internet of Things

**IL** Internal Link

**IT** Information Technology

**IDS** Intrusion Detection System

**MES** Manufacturing Execution System

**MQTT** Message Queueing Telemetry Transport

**MTBF** Mean Time Between Failure

**NFR** Non-functional Requirements

**NVD** National Vulnerability Database

**OEE** Overall Equipment Efficiency

**OPC UA** Open Platform Communication unified Architecture

**OT** Operational Technology

**OWL** Web Ontology Language

**PLC** Programmable Logic Controller

**P(N)** Probability of Node

**RCL** Role Class Library

**RTU** Remote Terminal Unit

**R(N)** Risk of Node

**SCADA** Supervisory Control and Data Acquisition

**SIEM** Security Information and Event Management

**S(N)** Severity of Node

**STAMP** System Theoretic Accident Model and Processes

**STPA** System Theoretic Process Analysis

**SUC** System Under Consideration

**SUCL** System Unit Class Library

**XML** Extensible Markup Language



# List of Figures

1.1	ICS general structure . . . . .	3
2.1	Risk management process . . . . .	21
3.1	Safety in industry . . . . .	26
3.2	Security in industry . . . . .	27
3.3	safety and security in industry . . . . .	28
3.4	Safety and security relationship: Antagonistic, Conditional Dependency, Mutual Reinforcement, and Independent. . . . .	29
4.1	DSRP modified for current research [100] . . . . .	31
4.2	Connection and components of the Sorting Station . . . . .	34
6.1	AML Simplified structure of Instance Hierarchy [33] . . . . .	46
6.2	AML model for the use case: Sorting Station (Fig. 4.2) . . . . .	48
6.3	System unit classes of Assets in ICS . . . . .	50
6.4	Instance hierarchy of Use Case . . . . .	51
6.5	Instance hierarchy, Internal Links, and Attributes example of Asset (PLC) . . . . .	52
6.6	Role Class Library for Sorting Station . . . . .	53
6.7	Expansion of Instance hierarchy with ILs . . . . .	54
6.8	Interface class for the sorting station . . . . .	55
6.9	Attributes for Assets of ICS . . . . .	57
6.10	AAS structured implementation [14] . . . . .	58
6.11	AAS model . . . . .	60
6.12	All AASs for the use case Sorting Station. . . . .	61
6.13	An example of PLC AAS along with submodels. . . . .	61
6.14	All submodels from PLC AASs for Sorting Station. . . . .	62
6.15	Ontology entities representation . . . . .	64
6.16	Ontology classes and their relations for Use Case . . . . .	65
6.17	Complete Ontology graphical representation using OntoGraph . . . . .	66
6.18	Representation of the Ontology in the Protégé in terms of (A) Classes, (B) Data properties, (C) Object properties, (D) Individual instances, and (E) Instance relation representation with other instances using object properties and data properties . . . . .	67
6.19	Reasoner result of the Ontology . . . . .	67

7.1	Concept of AML model development [19] . . . . .	72
7.2	AML model implementation . . . . .	72
7.3	Representation of (A) Node-to-node relationships and (B) Example CPT implementation for the BBN. . . . .	86
7.4	Representation of the BBN implementation. . . . .	92
8.1	Risk treatment procedure for integrated approach . . . . .	100
8.2	BBN risk propagation for use case without mitigation . . . . .	107
8.3	BBN risk propagation for use case with mitigation . . . . .	107

# List of Tables

2.1	Risk identification methods [109], [35]	13
2.2	Risk analysis and assessment methods	15
2.3	Risk assessment methods	18
5.1	Identified vulnerabilities from scan	41
5.2	Identified Inputs for Safety and Security Risk Assessment [132, 15]	43
6.1	Comparative analysis of information models as the single source of truth for risk assessment	70
7.1	CVSS Metrics for Probability and Severity Calculations	80
7.2	Calculated node probabilities P(N) and severities S(N).	83
7.3	Identified risks for the use case	84
7.4	Defined node-to-node relationships in the BBN model	85
8.1	vulnerability and its mitigations	102
8.2	Risk and its mitigations	102
8.3	Conditional Probability Table for Component Failure Based on Mitigations and Vulnerability	105
9.1	Calculated node risk probabilities R(N) for one year	111
9.2	Risk and its mitigations	113
9.3	Evaluation of the Risk Assessment Method Against ISO 31000 Principles	114



# List of Algorithms

7.1	Conditional Probability Assignment for Child Node Based on Vulnerability or Component Parent . . . . .	87
7.2	Conditional Probability Assignment for Child Node Based on Process or Hazard Parent . . . . .	88
7.3	Conditional Probability Assignment for Child Node Based on Component and Process Node . . . . .	89



# Bibliography

- [1] Threat landscape for industrial automation systems. statistics for h1 2021. *Kaspersky ICS CERT*, 09.09.2021.
- [2] Accident analysis of industrial automation. *BARPI - Bureau for analysis of industrial risks and pollution*, 2015.
- [3] Security risk assessment guide for industrial control systems. *Information-technology Promotion Agency, Japan Technology Headquarters IT Security Center (ISEC)*, April 2018.
- [4] Tschroub Abdelghani. Implementation of defense in depth strategy to secure industrial control system in critical infrastructures. *American Journal of Artificial Intelligence*, 3, 01 2020.
- [5] Shokoufeh Abrishami, Nima Khakzad, Seyed Mahmoud Hosseini, and Pieter van Gelder. Bn-slim: A bayesian network methodology for human reliability assessment based on success likelihood index method (slim). *Reliability Engineering & System Safety*, 2020.
- [6] Farman Afzal, Shao Yunfei, Mubasher Nazir, and Saad Mahmood. A review of artificial intelligence based risk assessment methods for capturing complexity-risk interdependencies: Cost overrun in construction projects. *Int. Journal of Managing Projects in Business*, 09 2019.
- [7] Jarmo Alanen, Joonas Linnosmaa, Timo Malm, Nikolaos Papakonstantinou, Toni Ahonen, Eetu Heikkilä, and Risto Tiusanen. Hybrid ontology for safety, security, and dependability risk assessments and security threat analysis (sta) method for industrial control systems. *Reliability Engineering & System Safety*, 220:108270, 2022.
- [8] Md. Tanjin Amin, Faisal Khan, Salim Ahmed, and Syed Imtiaz. A data-driven bayesian network learning method for process fault diagnosis. *Process Safety and Environmental Protection*, 150:110–122, 2021.
- [9] O.T. Arogundade, A. Abayomi-Alli, and S. Misra. An ontology-based security risk management model for information systems. In *Arab J Sci Eng* 45, 2020.

- [10] OVE Austrian Electrotechnical Association. Functional safety – safety instrumented systems for the process industry sector. part 1: Framework, definitions, system, hardware and application programming requirements. 2019-03-01.
- [11] Terje Aven. Risk assessment and risk management: Review of recent advances on their foundation. *European Journal of Operational Research*, 253, 2016.
- [12] Abdul Aziz, Salim Ahmed, and Faizal Khan. An ontology-based methodology for hazard identification and causation analysis. *Process Safety and Environmental Protection*, 123:87–98, 2019.
- [13] Alexander Belyaev and Christian Diedrich. Specification "demonstrator i4.0-language" v3.0. 07 2019.
- [14] Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. A centralised or distributed risk assessment using asset administration shell. In *2021 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–4, 2021.
- [15] Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. Automating safety and security risk assessment in industrial control systems: Challenges and constraints. In *2022 IEEE 27th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–4, 2022.
- [16] Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. Automationml use for safety and security risk assessment in industrial control systems. In *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–4, 2023.
- [17] Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. Integrated safety-security risk assessment for industrial control system: An ontology-based approach. In *2023 IEEE 28th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–8, 2023.
- [18] Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. Integrated safety-security risk assessment for production systems: A use case using bayesian belief networks. In *2023 IEEE 21st International Conference on Industrial Informatics (INDIN)*, pages 1–6, 2023.
- [19] Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. Automationml meets bayesian networks: A comprehensive safety-security risk assessment in industrial control systems. *IEEE Open Journal of the Industrial Electronics Society*, 5:823–835, 2024.
- [20] Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. Comparative analysis of aas and aml as a data source for integrated risk assessment in ics. In *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–4, 2024.

- [21] Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. Mapping ics vulnerabilities: Prioritization and risk propagation analysis with mitre att&ck framework and bayesian belief networks. In *2024 IEEE 29th International Conference on Emerging Technologies and Factory Automation (ETFA)*, pages 1–8, 2024.
- [22] Pushparaj Bhosale, Wolfgang Kastner, and Thilo Sauter. Modeling human error factors with security incidents in industrial control systems: A bayesian belief network approach. In *Proceedings of the 19th International Conference on Availability, Reliability and Security, ARES '24*, New York, NY, USA, 2024. Association for Computing Machinery.
- [23] Stefan Biffli, Luca Berardinelli, Emanuel Maetzler, Manuel Wimmer, Arndt Lueder, and Nicole Schmidt. Model-based risk assessment in multi-disciplinary systems engineering. In *2015 41st Euromicro Conference on Software Engineering and Advanced Applications*, pages 438–445, 2015.
- [24] Bernhard Brenner, Siegfried Hollerer, Pushparaj Bhosale, Thilo Sauter, Wolfgang Kastner, Joachim Fabini, and Tanja Zseby. Better safe than sorry: Risk management based on a safety-augmented network intrusion detection system. *IEEE Open Journal of the Industrial Electronics Society*, 4:287–303, 2023.
- [25] John F Carriger, Mace G Barron, and Michael C Newman. Bayesian networks improve causal environmental assessments for evidence-based policy. *Environmental science & technology*, 50(24):13195–13205, 2016.
- [26] Chao Chen, Genserik Reniers, and Nima Khakzad. Integrating safety and security resources to protect chemical industrial parks from man-made domino effects: A dynamic graph approach. *Reliability Engineering & System Safety*, 191:106470, 2019.
- [27] Qian Chen, Robert Abercrombie, and F.T. Sheldon. Risk assessment for industrial control systems quantifying availability using mean failure cost (mfc). *Journal of Artificial Intelligence and Soft Computing Research*, 5:205–220, 09 2015.
- [28] Sabarathinam Chockalingam, Dina Hadžiosmanović, Wolter Pieters, André Teixeira, and Pieter van Gelder. Integrated safety and security risk assessment methods: A survey of key characteristics and applications. In *Critical Information Infrastructures Security*, 2017.
- [29] Lawrence Chung, do Prado Leite, Julio Cesar, Sampaio Borgida, T. Alexander, Vinay K. Chaudhri, Paolo Giorgini, and Eric S. Yu. *On Non-Functional Requirements in Software Engineering, Conceptual Modeling: Foundations and Applications: Essays in Honor of John Mylopoulos*, pages 363–379. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

- [30] Brian Cohn, Todd Noel, Jeffrey Cardoni, Troy Haskin, Douglas Osborn, and Tunc Aldemir. Integrated safety and security analysis of nuclear power plants using dynamic event trees. *Nuclear Science and Engineering*, 197(sup1):S45–S56, 2023.
- [31] Jeff Cornelius. Why are industrial control system attacks increasing? *Magazine-InfoSecurity Group*, 22.09.2021.
- [32] Jürgen Dobaj, Christoph Schmittner, Michael Krisper, and Georg Macher. Towards integrated quantitative security and safety risk assessment. In *Computer Safety, Reliability, and Security*, 2019.
- [33] Rainer Drath. *AutomationML: A Practical Guide*. Walter de Gruyter GmbH, Berlin/München/Boston, 2021.
- [34] Rainer Drath. *AutomationML: The Industrial Cookbook*. Walter de Gruyter GmbH, Berlin/München/Boston, 2021.
- [35] Jesús Díaz-Verdejo, Abdelghafar M. Elhady, Hazem M. El-bakry, and Ahmed Abou Elfetouh. Comprehensive risk identification model for scada systems. *Security and Communication Networks, Hindawi*, 06.08.2019.
- [36] Matthias Eckhart, Bernhard Brenner, Andreas Ekelhart, and Edgar Weippl. Quantitative security risk assessment for industrial control systems: Research opportunities and challenges. *Journal of Internet Services and Information Security*, 9:52–73, 08 2019.
- [37] Matthias Eckhart, Andreas Ekelhart, Stefan Biffl, Arndt Luder, and Edgar Weippl. Qualsec: An automated quality-driven approach for security risk identification in cyber-physical production systems. *IEEE Transactions on Industrial Informatics*, PP:1–12, 01 2022.
- [38] Matthias Eckhart, Andreas Ekelhart, and Edgar Weippl. Automated security risk identification using automationml-based engineering data. *IEEE Transactions on Dependable and Secure Computing*, 10 2020.
- [39] Matthias Eckhart, Andreas Ekelhart, and Edgar Weippl. Automated security risk identification using automationml-based engineering data. *IEEE Transactions on Dependable and Secure Computing*, 19(3):1655–1672, 2022.
- [40] Marco Ehrlich, Andre Bröring, Dimitri Harder, Torben Auhagen-Meyer, Philip Kleen, Lukasz Wisniewski, Henning Trsek, and Jürgen Jasperneite. Alignment of safety and security risk assessments for modular production systems. *e & i Elektrotechnik und Informationstechnik*, 138:454–461, 15 September, 2021.
- [41] Sarah Ewing and Ronald Boring. Simulated human error probability and its application to dynamic human failure events. 10 2016.

- [42] Norman Fenton and Martin Neil. *Risk assessment and decision analysis with Bayesian networks*. Crc Press, 2018.
- [43] Ivo Friedberg, Kieran McLaughlin, Paul Smith, David Lavery, and Sakir Sezer. Stpa-safesec: Safety and security analysis for cyber-physical systems. *Journal of Information Security and Applications*, 34:183–196, 2017.
- [44] Thomas Gamer, Benjamin Kloepper, and Mario Hoernicke. The way toward autonomy in industry - taxonomy, process framework, enablers, and implications. In *IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society*, volume 1, pages 565–570, 2019.
- [45] Priscilla Grace George and V.R. Renjith. Evolution of safety and security risk assessment methodologies towards the use of bayesian networks in process industries. *Process Safety and Environmental Protection*, 149, 2021.
- [46] E. I. Gergely, D. C. Spoiala, V. Spoiala, H. M. Silaghi, and Z. T. Nagy. Design framework for risk mitigation in industrial plc control. In *2008 IEEE International Conference on Automation, Quality and Testing, Robotics*, volume 2, pages 198–202, 2008.
- [47] Matthias Glawe, Christopher Tebbe, Alexander Fay, and Karl-Heinz Niemann. Knowledge-based engineering of automation systems using ontologies and engineering data. In *Proceedings of the International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management, IC3K 2015*, page 291–300, 2015.
- [48] Rachael P.E. Gordon. The contribution of human factors to accidents in the offshore oil industry. *Reliability Engineering & System Safety*, 1998.
- [49] Lars Grunske, Robert Colvin, and Kirsten Winter. Probabilistic model-checking support for fmea. *Proceedings - 4th International Conference on the Quantitative Evaluation of Systems, QEST 2007*, pages 119–128, 10 2007.
- [50] Siegfried Hollerer, Marta Chabrová, Thilo Sauter, and Wolfgang Kastner. Combined modeling techniques for safety and security in industrial automation: A case study. In *2022 15th International Conference on Security of Information and Networks (SIN)*, 2022.
- [51] Siegfried Hollerer, Thilo Sauter, and Wolfgang Kastner. Risk assessments considering safety, security, and their interdependencies in ot environments. In *Proceedings of the 17th International Conference on Availability, Reliability and Security*, New York, NY, USA, 2022.
- [52] Ali M. Hosseini, Thilo Sauter, and Wolfgang Kastner. Safety and security requirements in aas integration: Use case demonstration. In *2023 IEEE 19th International Conference on Factory Communication Systems (WFCS)*, pages 1–8, 2023.

- [53] Kaixing Huang, Chunjie Zhou, Yu-Chu Tian, Weixun Tu, and Yuan Peng. Application of bayesian network to data-driven cyber-security risk assessment in scada networks. In *2017 27th International Telecommunication Networks and Applications Conference (ITNAC)*, pages 1–6, 2017.
- [54] Iec 61508: Functional safety of electrical/electronic/programmable electronic safety-related systems. *International Electrotechnical Commission (IEC)*., 2010.
- [55] Security for industrial automation and control systems –part 3-2: Security risk assessment for system design. *International Electrotechnical Commission*, 2020.
- [56] Industrial-process measurement, control and automation - framework for functional safety and security. *International Electrotechnical Commission (IEC)*, 2019.
- [57] Safety of machinery - security aspects related to functional safety of safety-related control systems. *International Electrotechnical Commission (IEC)*, 2019.
- [58] Vinay M. Ijure, Sean A. Laughter, and Ronald D. Williams. Security issues in scada networks. *Computers & Security*, 25, 2006.
- [59] Iso/iec 12100: General principles for design - risk assessment and risk reduction. *International Organization for Standardization (ISO)*, 2010.
- [60] Iso/iec 31000: Risk management - guidelines. *International Organization for Standardization (ISO)*, 2010.
- [61] Iso/iec guide 73: Risk management — vocabulary. *International Organization for Standardization (ISO)*, 2009.
- [62] Miguel A. Iñigo, Alain Porto, Blanca Kremer, Alain Perez, Felix Larrinaga, and Javier Cuenca. Towards an asset administration shell scenario: a use case for interoperability and standardization in industry 4.0. In *NOMS 2020 - 2020 IEEE/IFIP Network Operations and Management Symposium*, pages 1–6, 2020.
- [63] Juergen Jasperneite, Thilo Sauter, and Martin Wollschlaeger. Why we need automation models: Handling complexity in industry 4.0 and the internet of things. *IEEE Industrial Electronics Magazine*, 14(1):29–40, 2020.
- [64] Karl Johnson, Caroline Morais, and Edoardo Patelli. Ai tools for human reliability analysis. 07 2023.
- [65] L. Kaikkonen, T. Parviainen, M. Rahikainen, L. Uusitalo, and A. Lehikoinen. Bayesian networks in environmental risk assessment: A review. integrated environmental assessment and management. *Special Series: Applications of Bayesian Networks for Environmental Risk Assessment and Management*, 2020.
- [66] Peter E. Kaloroumakis and Michael J. Smith. Toward a knowledge graph of cybersecurity countermeasures. *The MITRE Corporation*, 2021.

- [67] Hiroo Kanamaru. Bridging functional safety and cyber security of sis/scs. In *2017 56th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*, 2017.
- [68] M. Karthick, C. Senthil Kumar, and T. Paul Robert. Bayes-hep: Bayesian belief networks for estimation of human error probability. *Life Cycle Reliability and Safety Engineering*, 2017.
- [69] Aram Kim, Junhyoung Oh, Kookheui Kwon, and Kyungho Lee. Consider the consequences: A risk assessment approach for industrial control systems. *Security and Communication Networks*, 2022.
- [70] C. Kolb, S. M. Nicoletti, M. Peppelman, and M. Stoelinga. Model-based safety and security co-analysis: a survey. *ArXiv*, abs/2106.06272, 2021.
- [71] Kevin B. Korb and Ann E. Nicholson. *Bayesian Artificial Intelligence*. CRC Press, 2010.
- [72] Andrew Kornecki, Subramanian N., and Zalewski Janusz. Studying interrelationships of safety and security for software assurance in cyber-physical systems: Approach based on bayesian belief networks. *2013 Federated Conference on Computer Science and Information Systems*, 01 2013.
- [73] Andrew Kornecki and Janusz Zalewski. Safety and security in industrial control. *ACM International Conference Proceeding Series*, 2010.
- [74] Andrew J. Kornecki, Nary Subramanian, and Janusz Zalewski. Studying interrelationships of safety and security for software assurance in cyber-physical systems: Approach based on bayesian belief networks. In *2013 Federated Conference on Computer Science and Information Systems*, 2013.
- [75] Daniel Krauß and Christoph Thomalla. Ontology-based detection of cyber-attacks to scada-systems in critical infrastructures. In *2016 Sixth International Conference on Digital Information and Communication Technology and its Applications (DICTAP)*, 2016.
- [76] Siwar Kriaa. *Joint Safety and Security Modeling for Risk Assessment in Cyber Physical Systems*. PhD thesis, 2016.
- [77] Siwar Kriaa, Marc Bouissou, Frederic Colin, Yoran Halgand, and Ludovic Pietre-Cambacedes. Safety and security interactions modeling using the bdmf formalism: Case study of a pipeline. In Andrea Bondavalli and Felicita Di Giandomenico, editors, *Computer Safety, Reliability, and Security*, pages 326–341, Cham, 2014. Springer International Publishing.
- [78] Siwar Kriaa, Marc Bouissou, Ludovic Piètre-Cambacedes, and Yoran Halgand. A survey of approaches combining safety and security for industrial control systems. *Reliability Engineering System Safety*, 139:156–178, 02 2015.

- [79] Siwar Kriaa, Marc Bouissou, and Ludovic Piètre-Cambacédès. Modeling the stuxnet attack with bdmp: Towards more formal risk assessments. In *2012 7th International Conference on Risks and Security of Internet and Systems (CRiSIS)*, pages 1–8, 2012.
- [80] D. Lee, J. Lee, Se-Woo Cheon, and J. Yoo. Application of system-theoretic process analysis to engineered safety features-component control system. 2013.
- [81] Seung Woo Lee, Ar Ryum Kim, Jun Su Ha, and Poong Hyun Seong. Development of a qualitative evaluation framework for performance shaping factors (psfs) in advanced mcr hra. *Annals of Nuclear Energy*, 38(8):1751–1759, 2011.
- [82] Tingting Li and Chris Hankin. A model-based approach to interdependency between safety and security in ics. *3rd International Symposium for ICS & SCADA Cyber Security Research 2015 (ICS-CSR 2015)*, 2015.
- [83] Mass Soldal Lund, Bjørnar Solhaug, and Ketil Stølen. *A Guided Tour of the CORAS Method*, pages 23–43. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.
- [84] Tyson Macaulay and Bryan L Singer. *Cybersecurity for Industrial Control Systems: SCADA, DCS, PLC, HMI, and SIS*. Auerbach Publications, London, 1 edition, 2012.
- [85] Y. A. Mahmood, A. Ahmadi, A. K. Verma, A. Srividya, and U. Kumar. Fuzzy fault tree analysis: a review of concept and application. *International Journal of System Assurance Engineering and Management*, 4, 2013.
- [86] James McCarthy, Michael Powell, Keith Stouffer, CheeYee Tang, Timothy Zimmerman, William Barker, Titilayo Ogunyale, Devin Wynne, and Johnathan Wiltberger. National institute of standards and technology interagency or internal (nistir) 8219 - securing manufacturing industrial control systems: Behavioral anomaly detection. *National Institute of Standards and Technology, U.S. Department of Commerce*, July 2020.
- [87] Martin Melik-Merkumians, Alois Zoitl, and Thomas Moser. Ontology-based fault diagnosis for industrial control applications. In *2010 IEEE 15th Conference on Emerging Technologies & Factory Automation*, 2010.
- [88] L. Mkrtchyan, L. Podofillini, and V.N. Dang. Bayesian belief networks for human reliability analysis: A review of applications and gaps. *Reliability Engineering & System Safety*, 139:1–16, 2015.
- [89] Mohammad Modarres and Se Woo Cheon. Function-centered modeling of engineering systems using the goal tree-success tree technique and functional primitives. *Reliability Engineering & System Safety*, 64(2):181–200, 1999.

- [90] Caroline Morais, Raphael Moura, Michael Beer, and Edoardo Patelli. Analysis and estimation of human errors from major accident investigation reports. *ASCE-ASME J Risk and Uncert in Engrg Sys Part B Mech Engrg*, 6, 10 2019.
- [91] Caroline Morais, Ka Lai Yung, Karl Johnson, Raphael Moura, Michael Beer, and Edoardo Patelli. Identification of human errors and influencing factors: A machine learning approach. *Safety Science*, 2022.
- [92] Bruno Augusti Mozzaquatro, Carlos Agostinho, Diogo Goncalves, João Martins, and Ricardo Jardim-Goncalves. An ontology-based cybersecurity framework for the internet of things. *Sensors*, 18(9), 2018.
- [93] Joseph Mukama. Master thesis - risk analysis as a security metric for industrial control systems. *Chalmers University of Technology, Department of Computer Science and Engineering*, Gothenburg, Sweden 2016.
- [94] Bård Myhre, Are Hellandsvik, and Stig Petersen. A responsibility-centered approach to defining levels of automation. *Journal of Physics: Conference Series*, 1357:012027, 10 2019.
- [95] N.R. Nayak, S. Kumar, and D. Gupta. Network mining techniques to analyze the risk of the occupational accident via bayesian network. *Int J Syst Assur Eng Manag* 13 (Suppl 1), 633–641, 2022.
- [96] Vivek Nigam, Alexander Pretschner, and Harald Ruess. Model-based safety and security engineering, 2018.
- [97] Valentina Di Pasquale, Salvatore Miranda, Walther Patrick Neumann, and Azin Setayesh. Human reliability in manual assembly systems: a systematic literature review. *IFAC-PapersOnLine*, 51(11):675–680, 2018. 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018.
- [98] M.Elisabeth Paté-Cornell. Uncertainties in risk analysis: Six levels of treatment. *Reliability Engineering & System Safety*, 54(2):95–111, 1996. Treatment of Aleatory and Epistemic Uncertainty.
- [99] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1988.
- [100] Ken Peffers, Tuure Tuunanen, Charles Gengler, Matti Rossi, Wendy Hui, Ville Virtanen, and Johanna Bragge. The design science research process: A model for producing and presenting information systems research. *Proceedings of First International Conference on Design Science Research in Information Systems and Technology DESRIST*, 02 2006.
- [101] Daniel Pereira, Celso Hirata, and Simin Nadjm-Tehrani. A stamp-based ontology approach to support safety and security analyses. *Journal of Information Security and Applications*, 47:302–319, 2019.

- [102] Kenneth Pettersen Gould and Corinne Bieder. *Safety and Security: The Challenges of Bringing Them Together*, pages 1–8. Springer International Publishing, 2020.
- [103] Algirde Pipikaite, Filipe Beato, and Georges De Moura. What the cyber-attack on the us oil and gas pipeline means and how to increase security. *World Economic forum*, 10 May, 2021.
- [104] Sandeep Pirbhulal, Vasileios Gkioulos, and Sokratis Katsikas. Towards integration of security and safety measures for critical infrastructures based on bayesian networks and graph theory: A systematic literature review. *Signals*, 2(4):771–802, 2021.
- [105] Pierre-Yves Piriou, Jean-Marc Faure, and Jean-Jacques Lesage. Generalized boolean logic driven markov processes: A powerful modeling framework for model-based safety analysis of dynamic repairable and reconfigurable systems. *Reliability Engineering and System Safety*, 163:57–68, 2017.
- [106] Ludovic Piètre-Cambacédès and Marc Bouissou. Modeling safety and security interdependencies with bdmp (boolean logic driven markov processes). In *2010 IEEE International Conference on Systems, Man and Cybernetics*, 2010.
- [107] Nayot Poolsappasit, Rinku Dewri, and Indrajit Ray. Dynamic security risk management using bayesian attack graphs. *IEEE Transactions on Dependable and Secure Computing*, 2012.
- [108] Joe Raad and Christophe Cruz. A survey on ontology evaluation methods. In *Proceedings of the 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management*, 2015.
- [109] Christian Raspotnig and Andreas Opdahl. Comparing risk identification techniques for safety and security requirements. *Journal of Systems and Software*, 2013. SI : Software Engineering in Brazil: Retrospective and Prospective Views.
- [110] Marvin Rausand and Stein Haugen. *Hazard Identification*, chapter 10, pages 259–337. John Wiley & Sons, Ltd, 2020.
- [111] Yeqi Ru, Yufei Wang, June Li, Jian Liu, Guotai Yang, Kai Yuan, and Kaipei Liu. Risk assessment of cyber attacks in ecps based on attack tree and ahp. In *2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, pages 465–470, 2016.
- [112] Alastair Ruddle, David Ward, Benjamin Weyl, Sabir Idrees, Yves Roudier, Michael Friedewald, Timo Leimbach, Andreas Fuchs, Sigrid Gürgens, Olaf Henniger, Roland Rieke, Matthias Ritscher, Henrik Broberg, Ludovic Apvrille, Renaud Pacalet, and Gabriel Pedroza. Security requirements for automotive on-board networks based on dark-side scenarios. 01 2009.

- [113] Enno Ruijters and Mariëlle Stoelinga. Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools. *Computer Science Review*, 15-16:29–62, 2015.
- [114] Sami Sader, István Husti, and Miklós Daróczy. Enhancing failure mode and effects analysis using auto machine learning: A case study of the agricultural machinery industry. *Processes*, 8(2), 2020.
- [115] Thilo Sauter, Stefan Soucek, Wolfgang Kastner, and Dietmar Dietrich. The evolution of factory and building automation. *IEEE Industrial Electronics Magazine*, 5(3):35–48, 2011.
- [116] Alejandro Seif, Carlos Toro, and Humza Akhtar. Implementing industry 4.0 asset administrative shells in mini factories. *Procedia Computer Science*, 159:495–504, 2019. Knowledge-Based and Intelligent Information & Engineering Systems: Proceedings of the 23rd International Conference KES2019.
- [117] Thomas Sheridan, W. Verplank, and T. Brooks. Human and computer control of undersea teleoperators. 12 1978.
- [118] Jyotsna Singh, Lakshay Mundeja, Armando W. Colombo, and Bilal Ahmad. Digitalization of the components of a mini factory with the implementation of rami 4.0 asset administration shell. In *7th International Conference on Industrial Cyber-Physical System*, pages 1–8, 2024.
- [119] Staab Steffen and Studer Rudi. In *Handbook on Ontologies*. Springer Berlin Heidelberg, 2009.
- [120] Mike StJohn-Green, R. Piggin, J.A. McDermid, and R. Oates. Combined security and safety risk assessment — what needs to be done for ics and the iot. In *10th IET System Safety and Cyber-Security Conference 2015*, 2015.
- [121] Keith Stouffer, Michael Pease, CheeYee Tang, Timothy Zimmerman, Victoria Pillitteri, Suzanne Lightman, Adam Hahn, Stephanie Saravia, Aslam Sherule, and Michael Thompson. NIST Special Publication 800-82 Revision 3: Guide to Operational Technology (OT) Security, September 2023.
- [122] Keith A. Stouffer, Victoria Pillitteri, Marshall Abrams, and Adam Hahn. Nist special publication 800-82 revision 2: Guide to industrial control systems (ics) security: Supervisory control and data acquisition (scada) systems, distributed control systems (dcs), and other control system configurations such as programmable logic controllers (plc). *National Institute of Standards & Technology*, 2015.
- [123] Weike Sun, Antonio R.C. Paiva, Peng Xu, Anantha Sundaram, and Richard D. Braatz. Fault detection and identification using bayesian recurrent neural networks. *Computers & Chemical Engineering*, 141:106991, 2020.

- [124] A. Sutcliffe, J. Galliers, and S. Minocha. Human errors and system requirements. In *Proceedings IEEE International Symposium on Requirements Engineering (Cat. No.PR00188)*, 1999.
- [125] Hayriye Tanyıldız, Canan Batur Şahin, and Özlem Batur Dinler. Improving deceptive patch solutions using novel deep learning-based time analysis model for industrial control systems. *Applied Sciences*, 14(20), 2024.
- [126] Christopher Tebbe, Karl-Heinz Niemann, and Alexander Fay. Ontology and life cycle of knowledge for ics security assessments. In *ICS-CSR '16: Proceedings of the 4th International Symposium for ICS & SCADA Cyber Security Research*, 2016.
- [127] Mahsa Teimourikia and Fugini Mariagrazia. Ontology development for run-time safety management methodology in smart work environments using ambient knowledge. *Future Generation Computer Systems*, 68:428–441, 2017.
- [128] Chee-Wooi Ten, Chen-Ching Liu, and Manimaran Govindarasu. Vulnerability assessment of cybersecurity for scada systems using attack trees. In *2007 IEEE Power Engineering Society General Meeting*, pages 1–8, 2007.
- [129] J. Anna Thangam, Subramania Bala Jeshurun, A. Thangapoo, S. Joe Patrick Gnanaraj, and M. Appadurai. Industrial hazards and safety measures – an empirical study. *Materials Today: Proceedings*, 2022. 4th Online International Conference on Science & Engineering of Material.
- [130] Dhanusha Tharanga. Critical review of risk identification techniques. *University of the West of Scotland*, 05 2020.
- [131] E. Theocharis, Michail Papoutsidakis, Christos Drosos, and G. Chamilothis. Safety standards in industrial applications: A requirement for fail-safe systems. *International Journal of Computer Applications*, 2019.
- [132] J. Tixier, G. Dusserre, O. Salvi, and D. Gaston. Review of 62 risk analysis methodologies of industrial plants. *Journal of Loss Prevention in the Process Industries 15 (2002) 291–300*, July 2002.
- [133] A. Treytl, T. Sauter, and C. Schwaiger. Security measures in automation systems-a practice-oriented approach. In *2005 IEEE Conference on Emerging Technologies and Factory Automation*, volume 2, pages 847–855, 2005.
- [134] P. Trucco, E. Cagno, F. Ruggeri, and O. Grande. A bayesian belief network modelling of organisational factors in risk analysis: A case study in maritime transportation. *Reliability Engineering & System Safety*, 2008.
- [135] Laura Uusitalo, Sakari Kuikka, Pirkko Kauppila, Pirkko Söderkultalahti, and Saara Bäck. Assessing the roles of environmental factors in coastal fish production in the northern baltic sea: A bayesian network application. *Integrated environmental assessment and management*, 8(3):445–455, 2012.

- [136] Siddhartha Verma, Thomas Gruber, Christoph Schmittner, and P. Puschner. Combined approach for safety and security. In *Computer Safety, Reliability, and Security*, 2019.
- [137] Li Wei, Yuan Ya-nan, Dong Wei-dong, and Wang Dan. Study on risk assessment of electric power construction project based on monte carlo simulation. In *2009 International Conference on Future BioMedical Information Engineering (FBIE)*, pages 532–535, 2009.
- [138] Shuang-Hua Yang Xiaorong Lyu, Yulong Ding. Safety and security risk assessment in cyber-physical system. *IET Cyber-Physical Systems: Theory & Applications*, 4-3:221–232, 2019.
- [139] Xuejiao Xing, Botao Zhong, Hanbin Luo, Heng Li, and Haitao Wu. Ontology for safety risk identification in metro construction. *Computers in Industry*, 109:14–30, 2019.
- [140] Xun Ye and Seung Ho Hong. Toward industry 4.0 components: Insights into and implementation of asset administration shells. *IEEE Industrial Electronics Magazine*, 13(1):13–25, 2019.
- [141] William Young and Nancy Leveson. Systems thinking for safety and security. In *Proceedings of the 29th Annual Computer Security Applications Conference, ACSAC '13*, page 1–8, 2013.
- [142] Esmaeil Zarei, Faisal Khan, and Rouzbeh Abbassi. Importance of human reliability in process operation: A critical analysis. *Reliability Engineering & System Safety*, 211:107607, 2021.
- [143] Hua Zhang, Fang Lou, Yunsheng Fu, and Zhihong Tian. A conditional probability computation method for vulnerability exploitation based on cvss. In *2017 IEEE Second International Conference on Data Science in Cyberspace (DSC)*, 2017.
- [144] Jinqing Zhang, Haosong Yue, Xingming Wu, and Weihai Chen. A brief review of bayesian belief network. In *2019 Chinese Control And Decision Conference (CCDC)*, pages 3910–3914, 2019.
- [145] Jinqing Zhang, Haosong Yue, Xingming Wu, and Weihai Chen. A brief review of bayesian belief network. In *2019 Chinese Control And Decision Conference (CCDC)*, 2019.
- [146] Qi Zhang, Chunjie Zhou, Naixue Xiong, Yuanqing Qin, Xuan Li, and Shuang Huang. Multimodel-based incident prediction and risk assessment in dynamic cybersecurity protection for industrial control systems. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 46(10):1429–1444, 2016.

- [147] Jiaqi Zhao, Matthias Schamp, Steven Hoedt, El-Houssaine Aghezzaf, and Johannes Cottyn. Automationml in industry 4.0 environment: A systematic literature review. In Philipp Weißgraeber, Frieder Heieck, and Clemens Ackermann, editors, *Advances in Automotive Production Technology – Theory and Application*, pages 162–169, Berlin, Heidelberg, 2021. Springer Berlin Heidelberg.