Method for Information and Process Modelling towards the Automation of Security Risk Assessments

Dissertation
zur Erlangung des akademischen Grades

Doktoringenieurin / Doktoringenieur
(Dr.-Ing.)

von M.Sc. Marco Ehrlich
geb. am 31.07.1989 in Höxter

genehmigt durch die Fakultät (Elektrotechnik und Informationstechnik/Maschinenbau/Verfahrens- und Systemtechnik) der Otto-von-Guericke-Universität Magdeburg

Gutachter/innen:
Prof. Dr.-Ing. Christian Diedrich
Prof. Dr.-Ing. Jürgen Jasperneite
Univ. Prof. Dipl.-Ing. Dr. techn. Wolfgang Kastner

Promotionskolloquium am 10.10.2024

Abstract

The importance of security for manufacturing systems is currently surging by the cause of two main aspects. Firstly, the increased modularity and flexibility of components, modules, and machines create a higher frequency of the required security risk assessments. Secondly, the highly dynamic threat landscape, the amount of available security-related information, and the growing degree of digitalisation emphasize the need for automated security risk assessments to support human experts with their currently mainly manual tasks. Therefore, this dissertation presents a method for information and process modelling towards the automation of security risk assessments for modular manufacturing systems. The first step of this method includes the information collection which uses swimlanes to specify a practical security risk assessment process based on the theoretical concepts from the available standardisation landscape. Afterwards, the second method step involves the information formalisation covering the integration of already established and acknowledged frameworks into the security risk assessment process in a technology- and implementation-agnostic way. Moreover, the third step of the method regarding information usage includes the elicitation of the human expert knowledge into rules based on predicate logic. Finally, the fourth method step includes the information access which describes the translation of the developed information model into a submodel of an asset administration shell, serving as the industrial implementation of a digital twin. By following these four method steps, the four main identified deficits of insufficient process coverage, missing standardised metrics, low approach maturity, and high abstraction levels within the state of the art are addressed. The results of this dissertation are the conceptual development and prototypical implementation of an expert system for automated security risk assessments regarding modular manufacturing systems. The overall evaluation shows a comparable result quality and process coverage between automated and manual performances of security risk assessments. Furthermore, the necessary level of knowledge for security risk assessments is decreased and the overall degree of automation is increased.

Kurzfassung

Die Bedeutung der Security für Fertigungssysteme nimmt derzeit aufgrund von zwei Hauptaspekten stark zu. Erstens steigt durch die zunehmende Modularität und Flexibilität von Komponenten, Modulen und Maschinen die Häufigkeit der erforderlichen Security Risikobewertungen. Zweitens unterstreichen die hochdynamische Bedrohungslandschaft, die Menge an verfügbaren sicherheitsrelevanten Informationen und der zunehmende Grad der Digitalisierung den Bedarf an automatisierten Security Risikobewertungen zur Unterstützung der menschlichen Experten bei ihren derzeit hauptsächlich manuellen Aufgaben. In dieser Dissertation wird daher eine Methode der Informations- und Prozessmodellierung zur Automatisierung von Security Risikobewertungen für modulare Fertigungssysteme vorgestellt. Der erste Schritt dieser Methode umfasst die Informationssammlung, die mit Hilfe von Swimlanes einen praktischen Prozess für Security Risikobewertungen auf der Grundlage der theoretischen Konzepte aus der verfügbaren Standardisierungslandschaft spezifiziert. Der zweite Methodenschritt beinhaltet die Informationsformalisierung, die bereits etablierte und anerkannte Rahmenwerke in den Prozess der Security Risikobewertung in einer technologie- und implementierungsunabhängigen Weise integriert. Darüber hinaus beinhaltet der dritte Schritt der Methode zur Informationsnutzung die Erhebung des menschlichen Expertenwissens in Form von Regeln, die auf Prädikatenlogik basieren. Der vierte Methodenschritt umfasst schließlich den Informationszugang, der die Übersetzung des entwickelten Informationsmodells in ein Teilmodell der Verwaltungsschale beschreibt, die als industrielle Umsetzung eines digitalen Zwillings genutzt wird. Mit diesen vier Methodenschritten werden die vier identifizierten Hauptdefizite des Standes der Technik - unzureichende Prozessabdeckung, fehlende standardisierte Metriken, geringe Reife der Ansätze und hohe Abstraktionsgrade - adressiert. Das Ergebnis dieser Dissertation ist die konzeptionelle Entwicklung und prototypische Implementierung eines Expertensystems zur automatisierten Security Risikobewertung von modularen Fertigungssystemen. Die Evaluierung zeigt eine vergleichbare Ergebnisqualität und Prozessabdeckung zwischen automatisierter und manueller Durchführung von Security Risikobewertungen. Darüber hinaus wird der notwendige Kenntnisstand möglicher Nutzer gesenkt und der Automatisierungsgrad mit Bezug auf Security Risikobewertungen insgesamt erhöht.

Contents

Part I.

Prologue

1. Introduction

1.1. Motivation

International and globalised markets are rapidly changing due to adapted customer requirements, smaller product volumes, shorter lifecycles, volatile prices, and expanding technological possibilities, especially in the area of information processing [1]. Hence, manufacturing systems are facing a growing amount of challenges regarding processes, people, and technologies [2, 3]. An adequate response lies in the overall developments regarding **Industrie 4.0 (I4.0)** and the **Industrial Internet (II)** [4, 5]. There, a lot of efforts are put into the increased digitalisation by integrating Information Technology (IT) into Operational Technology (OT) environments, flexibility, modularisation, and networked automation to enhance the overall capabilities of manufacturing systems [1, 6–10]. This enables a disruptive paradigm change within the manufacturing domain leading to a complex interconnection of physical and digital realms and a break-up of the traditional automation pyramid [11, 12].

In this context, high-level specifications for future manufacturing systems foresee a hybrid landscape of components and networks containing pervasive wired and wireless solutions intertwined, often comprising legacy or isolated systems [13–16]. Simultaneously, improvements concerning the overall engineering, monitoring, and performance of heterogeneous networking infrastructures are required in a reliable, secure, and automated way [17, 18]. Nevertheless, the current situation inside the OT domain differs from these described visions due to the typical brownfield architectures, which have been developed in a highly specialised manner and are dedicated to particular applications with specific requirements, such as determinism, a high availability, and long system lifetimes [19, 20]. This prevalent heterogeneity results in increased efforts, time, and resources required for the typical tasks regarding the management of **Fault, Configuration, Accounting, Performance and Security (FCAPS)**[1]. In consequence, the gap between the rising demands and the increased necessity of the prevalent manual tasks for the available security experts is widening [21–24]. In order to reduce this performance gap, the formalisation of information models needs to be integrated into software tools to automate the currently manual engineering steps to an acceptable level for operators [17, 25–27].

This highly dynamic mixture of systematic, organisational, and technological advances also needs to be investigated from the engineering viewpoints of **safety** (protection of humans, machines, and environment) as well as **security** (protection of machines from human adversaries) [28, 29]. Only thereby, a safe and secure operation

[1] https://www.itu.int/rec/T-REC-M.3400-200002-I/en

of future manufacturing systems without impacts on the availability, productivity, or personnel can be guaranteed (as also required by the regulative frameworks further discussed within Section 3.3). This robustness is also becoming a characteristic of quality for systems and creates incentives for implementing the associated measures. Hence, safety and security are obligatory factors for overall success, acting as enablers for future developments [30–32]. In addition, safety is defined as an essential function within the proposed suite of security standards[2] by the International Society of Automation (ISA) and the International Electrotechnical Commission (IEC), abbreviated as IEC 62443 within the rest of this dissertation. Furthermore, the manual and singular processes of testing, assessing, and certifying safety and security are changing towards continuous, technology-supported, automated, and information-driven processes [33]. Following these adaptations, the required results will be cheaper to obtain, faster to achieve, traceable, reproducible, and with a comparable quality level [33]. Therefore, this dissertation will focus on the automation of **security engineering** during operation, especially the intrinsically mandatory process of **security risk management** needed for Governance, Risk Management, and Compliance (GRC) requirements. For manufacturing systems, this is typically implemented via a **security risk assessment** process including risk identification, risk analysis, and risk evaluation. Figure 1.1 illustrates the embedding of the involved topics and the clarification of the described focus in grey. Afterwards within the next two sections, the other aspects of the overall motivation are presented within the application scenario and the problem statement.

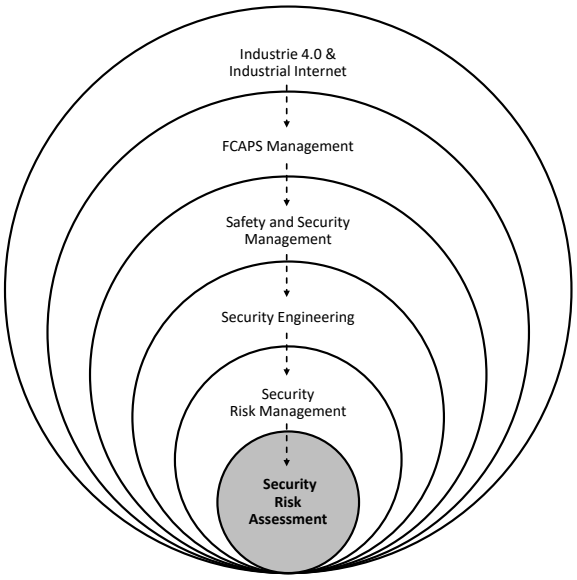

Figure 1.1: Overview and alignment of relevant topics for this dissertation

[2]https://www.isa.org/standards-and-publications/isa-standards/isa-iec-62443-series-of-standards

1.2. Application Scenario

The development and research of agile manufacturing systems for industrial production are ongoing topics since the 1990s, mainly driven by market pressure based on competitiveness and price considerations [34, 35]. In addition, the widening customer expectations, the demand for highly customised products with sufficient quality, the synthesis of diverse technologies, and the needed capabilities for reconfiguration lead to the four general types of manufacturing systems we know today [3]:

1. Dedicated Manufacturing Systems (DMSs)
2. Flexible Manufacturing Systems (FMSs)
3. Adjustable Manufacturing Systems (AMSs)
4. Reconfigurable Manufacturing Systems (RMSs)

Especially the reconfigurability of agile manufacturing systems is one of the main drivers for the I4.0 domain to provide factories with the needed adaptability [36]. The RMSs with their corresponding requirements regarding security from the viewpoint of system integrators and asset owners are one of the main aspects of motivation for this dissertation. RMSs can be generally described by six core characteristics [3]:

1. Modularity: The systematic breakdown of functionality into smaller parts that are designed and produced independently of one another but, when combined, work together seamlessly.
2. Flexibility: The capability of modifying the hardware and software configuration by adding, removing, or changing components.
3. Interoperability: The capability of exchanging information and resources, syntactic and semantic, among devices from different vendors.
4. Integrability: The ability to integrate new modules into the system without disruptions.
5. Referenceability: The capability to provide an effective method to document all changes in a system for future reference and provide a means for collaboration.
6. Comprehensibility: The capability to help enhance an individual's understanding of the system by focusing on clarity, transparency, and traceability.

The motivating application scenario as the main use case for this dissertation is defined as **Security Risk Assessment for Modular Manufacturing Systems** and is used to analyse requirements, to identify the existing problems, and to propose the associated solution approach. One key feature of RMSs is the capability to add or remove individual components, such as sensors, actuators, or controllers, or

complete modules, e.g. complex machinery or production cells, in a rapid and easy manner due to reduced engineering and configuration efforts [3, 6]. This creates modular manufacturing systems that are able to adapt to new product or process requirements [26, 37]. Figure 1.2 shows a conceptual view of the application scenario based on a generic RMS. The running and functioning modular manufacturing system consists of several (1...m) static modules plus additional network infrastructure. Depending on the process or product requirements, various flexible modules (1...n) can be added or removed to the modular manufacturing system to adapt to certain needed functionalities, such as laser engraving, drilling, or quality assurance.

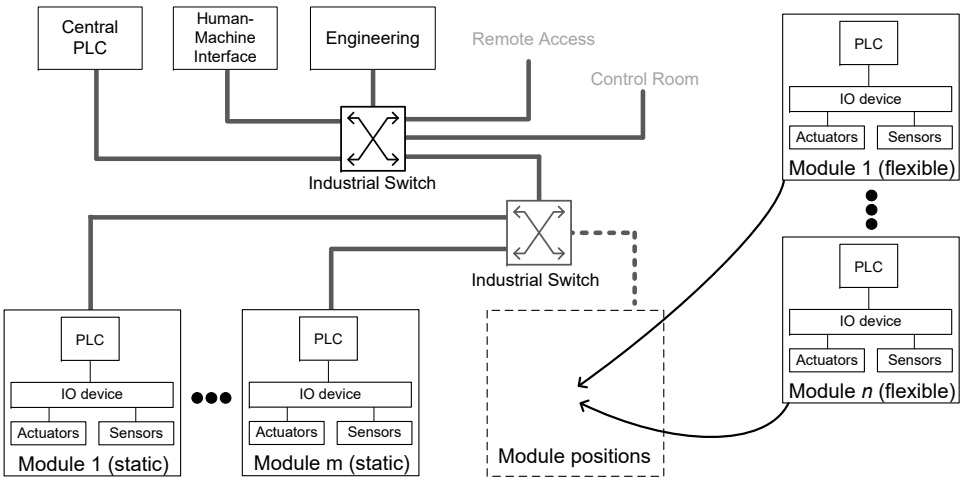

Figure 1.2: Conceptual view on the application scenario (adapted from [38])

A real-world example of the described application scenario can be found within the SmartFactoryOWL[3], a joint institution of the Fraunhofer IOSB-INA[4] and the OWL University of Applied Sciences and Arts (TH OWL)[5] in Lemgo, Germany. The respective industrial-grade demonstrator from the research environment is named Customisable Production System and is able to produce small LEGO figurines in a modular production process. It contains several different modules, e.g. a collaborative robot, a laser engraving cell, a handcraft working space, and an optical quality assurance, which are each connected via a set of conveyor belts. This represents a setup of a modular manufacturing system including technologies from various manufacturers and different communication and control systems [39]. Therefore, it is used as a typical modular Industrial Automation and Control System (IACS) as the System Under Consideration (SUC) (as defined within the IEC 62443-1-1 standard) of this dissertation. Additional information about the utilised industrial demonstrator can be found within Section 5.1.

[3]https://smartfactory-owl.de/?lang=en
[4]https://www.iosb-ina.fraunhofer.de/en.html
[5]https://www.th-owl.de/en

1.3. Problem Statement

In addition to the mentioned aspects from the application scenario, there is also an engineering perspective being part of the overall motivation. Every modular manufacturing system needs a continuous certification based on a safety risk assessment to be operated and to comply with all the regulative requirements imposed by present laws [40]. These general needs are recorded within the Machinery Directive from 2006, respectively the 9th ProdSG (Product Safety Act) in Germany, and just recently within the Machinery Regulation from 2023 for the European markets and derived into safety standards, such as the IEC 61508 and the IEC 61511 [6, 41]. The interrelated security risk assessments are mandatory as well, showing the coupling of both domains. This trend within the security domain began with critical infrastructures and is currently expanding into further industrial domains that are imposed with a necessity for security [42]. In addition, these safety risk assessments are required for first-time commissioning and after every major functional change of an agile manufacturing system. In detail, a change can be internal, e.g. due to a component exchange, an update of an asset via patches, or a replacement of a complete module, or external, e.g. due to new regulations or standards, newly discovered vulnerabilities, newly detected threat actors or techniques, and known incidents. Especially when the security of a safety function is affected by a change, the security risk assessment is mandatory regarding the definitions from the IEC 61508 [43]. So far, it has been sufficient to perform a security risk assessment during the commissioning of a manufacturing system once, as subsequent reconfigurations were typically either minor or infrequent [31]. Nevertheless, flexible production is an essential requirement for the I4.0 developments and implies a high level of modularity and a high frequency of reconfigurations. Today, it is necessary to reconfigure the modular manufacturing systems not only during the course of scheduled maintenance times, but even amid regular operation for maximum efficiency [3].

Currently, the security of a modular manufacturing system has to be re-evaluated manually after every reconfiguration by security experts, including an analysis of the implemented configuration, a security risk assessment, and the corresponding documentation [44]. Thereby, it means that security risk assessments have to be performed as a routine task in a continuous, iterative, and cyclic manner with a high frequency [45]. This imposes high efforts with regard to time, financial cost, and general resources and prevents scalability due to the necessity of error-prone manual tasks leading to a trade-off between the dynamics of flexible production and its required security [3, 46, 47]. In addition, today's relevant standards for security risk assessments were originally designed for general static manufacturing systems in isolated environments and now need to be applied to specific modular manufacturing systems within interconnected architectures [45]. Furthermore, the frequency of security risk assessments needs to rise as well due to the surge of recent security incidents, publicly available information about technical vulnerabilities, the intensifying threat landscape, and the increasing extent of possible impacts [48]. Figure 1.3 visualises the resulting performance gap as one of the key problems for

this dissertation and the expected improvement that can be achieved using the proposed solution approach described. Security risk assessments have been the subject of research within recent years to develop new approaches to reduce the efforts for a safe and secure operation of modular manufacturing systems [3]. This is mainly due to the lack of reusable and shareable knowledge, automated processes, process coverage, and the measurement of effective security [49]. The overall aim is always to reduce the efforts during operation, because typically the Operational Expenditures (OPEXs) make up to 80% of the total costs in comparison to the Capital Expenditures (CAPEXs) [40].

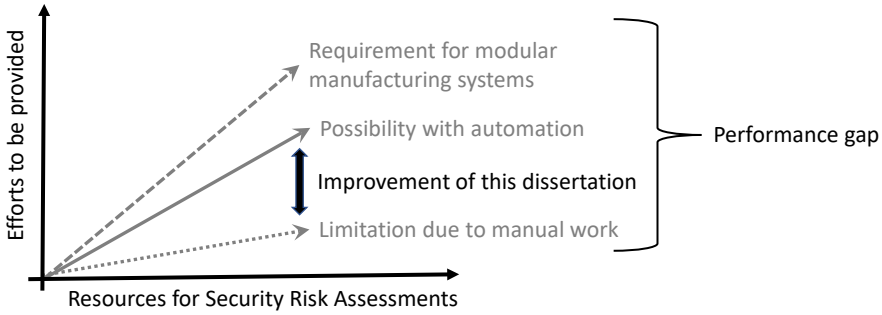

Figure 1.3: Visualisation of the performance gap regarding modular manufacturing systems and the associated security risk assessments (adapted from [50])

The stated problem exposition is relevant for every kind of business independent of size which is operating modular manufacturing systems for their production. However, this proves to be true especially for Small and Medium-sized Enterprises (SMEs) due to a plethora of security-related challenges, such as missing support of all security risk assessment stages, a low degree of automation, the lack of security and safety coupling, or proprietary data sources and information models. A complete list of challenges in scope for this dissertation regarding people, processes, and technology can be found within Table 3.4 in Section 3.5.

The current status of security risk assessments regarding modular manufacturing systems and their utilisation within businesses impair technological advances from other engineering domains as well, e.g. automatic configuration, self-X concepts, plug and produce, and the implementation of Machine Learning (ML) or Artificial Intelligence (AI) algorithms. In general, there is a lack of research regarding the topics of measuring, assessing, verifying, testing, sharing, and automating security [49]. Following the NA/NE 35 document provided by the German NAMUR association, security and safety are essential parts of the three main phases of Conceptual Engineering, Basic Engineering, and Detail Engineering which sum up to around 90% of the total direct engineering efforts based on the related example of systems from the process industry domain. Furthermore, in [44] the topic of OT security is described as an important aspect for every engineering phase of a system from the requirements analysis until decommissioning. Therefore, automated security risk

assessments are needed to make faster decisions regarding production security, to improve and streamline the quality of results, and to increase the overall degree of automation [51]. This dissertation aims to narrow the gap between the ever-increasing complexity of modular manufacturing systems and their security to reduce the work of the operators on site to an acceptable amount of resources and efforts.

The overall analysis of the available related work (see Section 4.3.3) reveals certain additional research gaps. These are defined here as the four main deficits (D1 - D4) regarding the automation of security risk assessments for modular manufacturing systems which will be addressed within this dissertation:

- **D1: Insufficient process coverage** - The complete security risk assessment process is not covered, e.g. risk analysis and risk evaluation are missing which results in a lack of final risk determination. In addition, the coupling with safety processes, models, or objectives is not regarded within the related work.
- **D2: No standardised metrics** - Typically, proprietary formalisations are proposed that contrast well-established security metrics and block the reusage of knowledge. Moreover, the general conformity towards globally-accepted standards is missing. A lack of details and poor user guidances hinders the understanding of the approaches. In addition, the usage of subjective data limits the comprehensibility.
- **D3: Low maturity** - There are rarely implemented and evaluated approaches available. The maximum Technology Readiness Level (TRL) is typically 4. Furthermore, the improvement of the automation degree is often not clear in terms of resource savings or knowledge formalisation and the security improvement is therefore not measurable. In many cases a sophisticated Graphical User Interface (GUI) as interface and improvement of the usability for the operator is missing as well.
- **D4: High abstraction level** - The proposed models are typically not suitable for real-world scenarios due to missing interfaces towards assets or a generic concept level. Furthermore, most of the models are not directly applicable and need additional manual refinement for the targeted use cases.

The possible gains and improvements can be estimated in a twofold way. Firstly, the company HIMA states that recurring tasks for functional safety should be defined once and their implementation in terms of time and content should then be automatically monitored. Specialist personnel can thus concentrate on important process steps, which in practice results in cost savings in the range of 70%[6]. In addition, the company Merck Electronics calculated the possible gains of a modular manufacturing system in comparison to a conventional system via a practical example from the process industry domain. The results show a possible time reduction of

[6]https://www.prozesstechnik.industrie.de/aufmacher/hima-stellt-digitalisierungsstrategie-vor

about 50% and a cost reduction up to 30% [52]. Secondly, an analysis of the specific characteristics of security risk assessments regarding time and financial costs has been done using typical industrial-grade demonstrators (see Section 5.1) to get a baseline measurement for the needed resources [53]. These two approaches lead to the following deduction of two main hypotheses (H1 and H2) supporting the assumption that an automation of security risk assessment processes can achieve complete, quality-steady, and acceptable results and a reduction of the needed efforts for modular IACSs. Therefore, an automation of security risk assessments for the industrial domain is generally of great interest [54, 55].

> *H1 (Result Quality): If security risk assessments for IACSs are automated, the results are qualitatively comparable to their manual counterpart.*

> *H2 (Process Automation): If the automation degree of security risk assessments for IACSs is increased, the overall required efforts for the operator are reduced.*

1.4. Solution Approach

In order to address the previously identified problems, this dissertation aims at designing a semi-formal and self-contained information model as the basis of an expert system covering security risk assessments for modular IACSs, which is shown in Figure 1.4 [56, 57]. Furthermore, an associated implementation utilising Digital Twins (DTs) is developed to transfer the information model into a usable and automated software prototype [32]. In doing so, it is possible to develop more practical results in comparison to the related works that mainly stay on the conceptual level. This practicability is achieved via an exchange of flexibility and inference capabilities towards a high credibility and reproducibility of security risk assessment results. The achieved information model and expert system of this dissertation are designated for system integrators and asset owners (as defined within the IEC 62443-3-2 security standard) during the development, commissioning, operation, and maintenance lifecycle phases (as defined within the Reference Architecture Model Industrie 4.0 (RAMI4.0)) for modular systems (as defined within the Verein deutscher Ingenieure (VDI) 2776-1 standard). Several aspects of the solution approach are provided as integral parts to the AutoS^2 research project[7].

The information model is used to establish a common understanding and vocabulary about the security risk assessment processes including the necessary knowledge, decisions, and data formats. Additionally, it abstracts the inherent complex characteristics in a technology- and implementation-independent way by defining assumptions

[7] https://www.init-owl.de/en/research/projects/detail/automatische-bewertung-und-ueberwachung-von-safety-security-eigenschaften-fuer-intelligente-technische-systeme

and surrounding conditions, specifying used and unused facts, determining the underlying process, and including the allowed inputs, outputs, internal procedures, rules, and operations. Typical standards in this domain, e.g. the IEC 62443, VDI/Verband der Elektrotechnik Elektronik Informationstechnik (VDE) 2182, or the International Organization for Standardization (ISO)/IEC 27000, are specified horizontally on a general level. Therefore, this dissertation proposes a specific information model as a vertical puncture for the automation of security risk assessments. To achieve this, it is expressed in a graphical language plus the corresponding syntax and semantic. This enables elicitation and reusability of expert knowledge to create traceable and reproducible results with comparable completeness, metrics, and quality of security risk assessments for modular manufacturing systems.

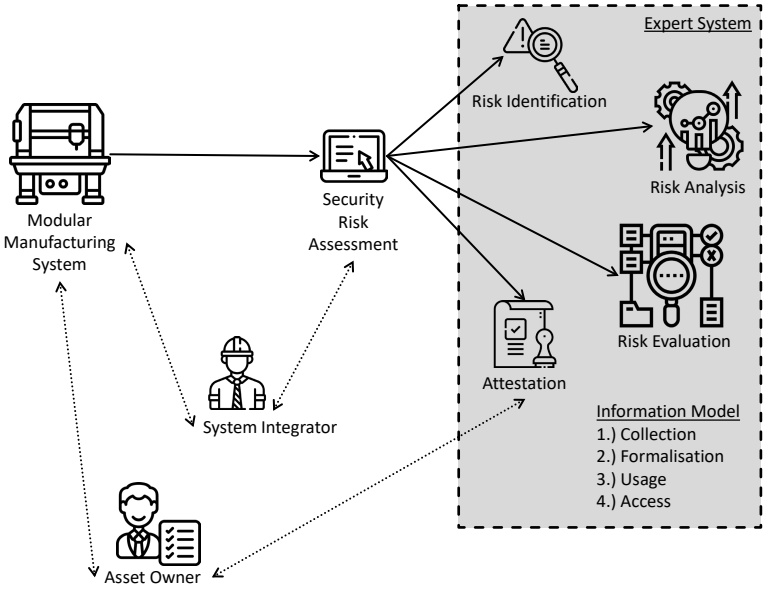

Figure 1.4: Overview of the proposed solution approach of this dissertation

The approach for the design of the information model consists of four main method steps which are adapted from the general knowledge engineering process defined in [58] and can be seen as the guiding research questions for this dissertation [56, 57].

1. How to collect the general information needed for an automated security risk assessment for modular IACSs? (Information Collection)

2. How to formalise a semi-formal information model for an automated security risk assessment? (Information Formalisation)

3. How to automate the decisions necessary within a typical manual security risk assessment in an informed way? (Information Usage)

4. How to integrate the information model into a DT for an automated security risk assessment? (Information Access)

Information Collection describes the first step for the creation of the information model. It includes the analysis of currently existing processes, technologies, and concepts to identify the required information for complete, high-quality, and practically usable security risk assessments. Furthermore, this includes the evaluation of available standardisation and certification processes to provide a detailed view on the required tasks in relation to, e.g. needed domain-specific knowledge, amount of required efforts and resources, degree of automation, involved stakeholders, and the coupling of safety and security. The focus is set on the IEC 62443 as the currently most important security standard for IACSs, which also includes modular manufacturing systems per definition, due to its novelty, popularity, and usage on a global scale. The subpart IEC 62443-3-2 covers the topic of security risk assessments based on a variety of Zone and Conduit Requirements (ZCRs) representing the individual process steps. Nevertheless, the standard describes the inherent steps and the corresponding information only on a very abstract level and the necessary details for a concrete information model and a practical implementation are omitted [56]. The achieved main results are:

- Requirements analysis for the information collection → Section 6.1
- Security risk assessment definition visualised via swimlanes → Section 6.2

Information Formalisation has the aim of specifying the abstraction of necessarily required information for security risk assessments into a technology- and implementation-independent way [59]. In addition, already established and acknowledged approaches, such as asset characteristics described with pre-existing ECLASS[8] properties, threats and mitigations specified via the MITRE Industrial Control System (ICS) ATT&CK[9] framework, or vulnerabilities represented through Common Vulnerability Enumeration (CVE)[10] entries, are integrated to establish a high credibility and adaptability. By formalising the information model in the described way, it inherits a sufficient level of abstraction and hierarchy. In addition, an object-oriented and interoperable semi-formal information model as a basis for the automation of security risk assessments is achieved which specifies the set of required and allowed inputs, the internal procedures, rules, and operations, and the set of possible outputs [56]. In contrast to formal models which are typically provable mathematically, the semi-formal information model of this dissertation is used to establish a common understanding and communication of the contents with a syntax and a semantic to enable the elicitation and reusability of expert knowledge [59]. The achieved main results are:

- Requirements analysis for the information formalisation → Section 7.1
- Unified Modelling Language (UML) class diagrams → Section 7.2

[8] https://www.eclass.eu/en/eclass-standard/search-content
[9] https://www.attack.mitre.org/tactics/ics
[10] https://www.cve.org

Information Usage includes the analysis of possible approaches to utilise the designed information model, to integrate the collected expert knowledge, and to enable automated decision making processes. Currently, there are no data sets publicly available that describe the inherited processes of security risk assessments. Therefore, all learning-based approaches from the domains of ML and AI for the automation of processes cannot be used. To overcome this scarcity and the need for available training data, the research concepts of static decision trees and other logical structures were analysed to use the identified, collected, and formalised information. The chosen predicate logic enables the definition of general expert knowledge and the specification of decision rules which are needed to perform automated security risk assessments for modular manufacturing systems [56]. In addition, the results of the automated security risk assessments need to have a clear transparency and a high credibility to create trustworthiness regarding the users. The achieved main results are:

- Requirements analysis for the information usage → Section 8.1
- Expert knowledge elicitation based on predicate logic rules → Section 8.2

Information Access describes the process of translating the complete information model into an acknowledged and machine-readable format. This is essentially required for an acceptable prototypical implementation and the provisioning of predefined libraries for reusability. The already standardised Asset Administration Shell (AAS) as the industrial implementation of a DT and the associated data structure is used to integrate the information model for security risk assessments as a new proposal for a submodel [53]. By these means, the elaborated information model representation can be used in an interoperable way with already present approaches and does not influence normal operation. For clarification, this dissertation focuses on performing security risk assessments of modular IACSs that are represented by AASs. The aim is not to increase the security of the AAS itself. The achieved main results are:

- Requirements analysis for the information access → Section 9.1
- Information model translation into the AAS data structure → Chapter 9.3

Figure 1.5 shows the overall dissertation structure in a chronological and procedural order based on the four method steps of information collection, information formalisation, information usage, and information access. The final result is an expert system covering the security risk assessment process in an automated manner. This figure is also used as a repetitive guidance throughout the complete document within the intermediate summaries of each distinctive method step. The further scoping of this dissertation is presented within Section 5.

The overall result of this dissertation is a knowledge-based forward-chained rule-based expert system with semi-formal, qualitative, and deterministic characteristics, which is further described within Section 4.2. The definition of an expert systems

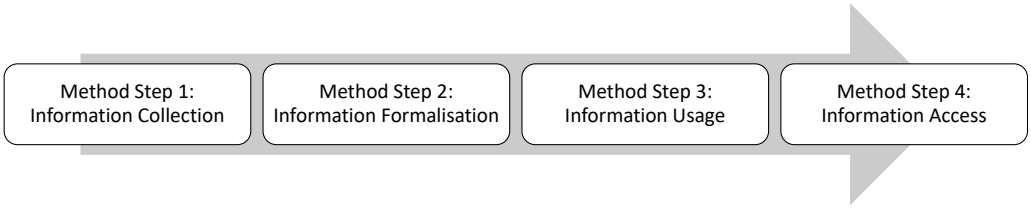

Figure 1.5: Dissertation structure based on the four methods steps used for repetitive guidance through the document

is as follows (translated from German [60]): "An expert system should be thought of as a program in which the competence of experts, who are excellently versed in a narrowly limited area, is bundled in a knowledge base and made available in an information technical adequate form". This expert system is able to acquire and to structure knowledge, to understand and to solve the problem with a high reliability, and to explain and evaluate the security risk assessment results [60]. The known and formalised knowledge builds up the foundation for the data-driven result creation process, which is called forward-chaining. By doing so, unknown security risk assessment results can be concluded from known starting conditions. This style is highly recommended to be used for OT security analyses within SMEs [44].

The domain of ML or AI approaches is not further regarded due to various reasons [61]. Objective, accurate, and complete training data sets are missing to create a fundamental background knowledge to feed learning-based algorithms [62–65]. Safety and security risk assessment processes need to be based on a high level of credibility and reproducibility, therefore black box systems that do not offer explanations with regard to their internal procedures will not be trusted by industrial stakeholders [63, 66]. Furthermore, security-related attacker behaviour and characteristics typically do not follow any probabilities in a quantitative way [65]. The implementation of an automated risk assessment will perform checks on high-stakes or safety-critical applications which have an increased level of impact [58]. Therefore, adequate technical and ethical standards for ML or AI approaches need to be present in the future [58].

To achieve the described objectives, a defined research method was developed and pursued. The complete procedure is shown in Figure 1.6 summarising the chosen research method of this dissertation which is based on the chronological performance of various research aspects. In addition, the achieved deliverables and publications are displayed to get an overview of the whole working scope. This is done via the names of the associated conferences or documents and the corresponding year of publication. In addition, the current publications which are already submitted but still under review are shown in light grey. The complete list of relevant publications for this dissertation can also be found within Annex D in Section 17.

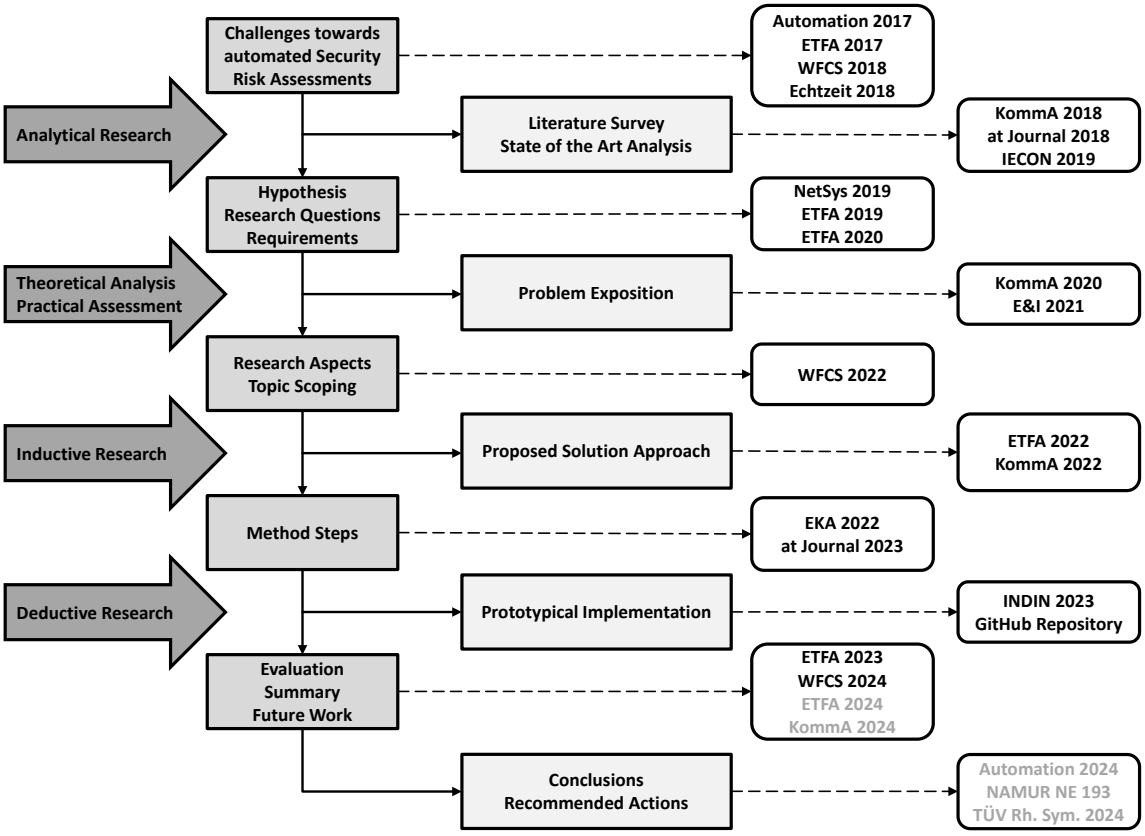

Figure 1.6: Research method including deliverables and publications throughout the course of this dissertation (adapted from [67])

2. Outline

The remainder of this dissertation is structured as follows and summarised in Figure 2.1 below. In Section 3 the fundamental background as the relevant base of contents is presented. The current state of the art and the related work is discussed within Section 4 to identify the main deficits regarding the automation of security risk assessments. Section 5 summarises the focus and the scoping of the method steps for the information and process modelling. In Section 6 the method step of information collection is described based on the definition of swimlanes. Afterwards, Section 7 describes the method step of information formalisation by using UML class diagrams for modelling. Section 8 presents the method step of information usage by providing the expert knowledge elicitation in the form of logic-based rules. Furthermore, Section 9 discusses the method step of information access containing the translation of the defined information model into a DT. The next part of this dissertation includes the overall result discussion based on the description of the prototypical implementation in Section 10 and the evaluation including the aspects of verification, validation, and credibility in Section 11. Finally, Section 12 concludes this dissertation and Section 13 points out possible future work.

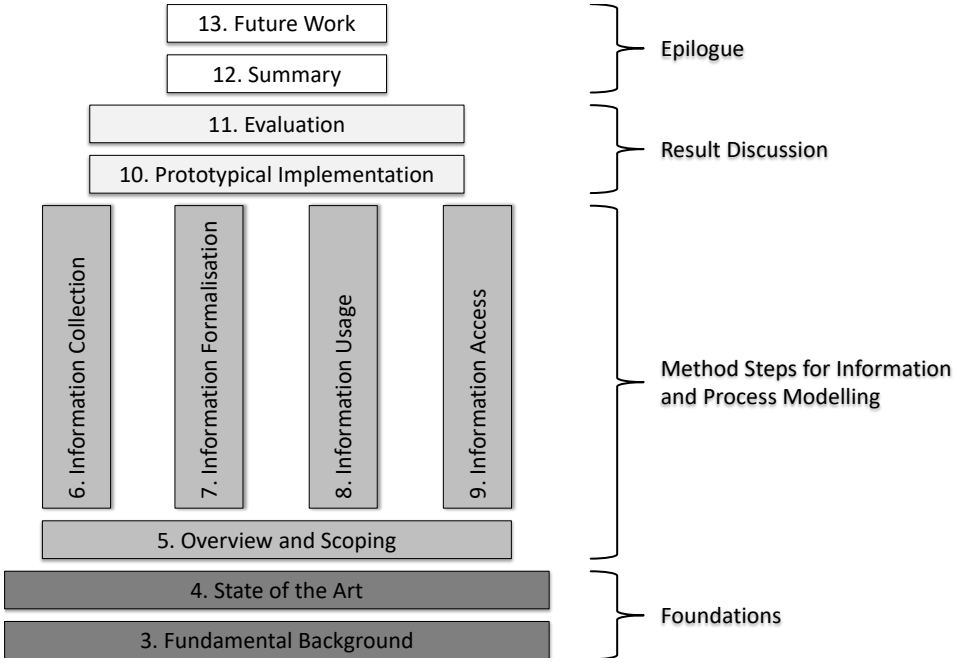

Figure 2.1: Structure of this dissertation regarding the following sections

Part II.

Foundations

3. Fundamental Background

In the following, Section 3.1 provides a definition for security used within this dissertation as well as an overview of associated concepts and involved stakeholders. Section 3.2 describes the ongoing IT/OT convergence including a comparison of the corresponding architectures, domain characteristics, security objectives, and a first alignment towards safety. Afterwards in Section 3.3, a chronological summary of the associated legal requirements and frameworks is displayed to emphasise the increasing need for automated security risk assessments. Furthermore, the coupling of security and safety is presented and discussed in Section 3.4. In Section 3.5, the external and environmental aspects surrounding this dissertation are illustrated based on security-related incidents, surveys, and challenges to sharpen the understanding and motivating the importance of the proposed solution approach. At the end, Section 3.6 describes the process covered in this dissertation and defines security risks as well as security risk assessments within the scope of the IEC 62443 standard.

3.1. Security Definition and Scoping

As highlighted in the introduction, the focus of this dissertation is set on the automation of security risk assessments for modular IACSs. In general, security is a non-functional requirement and a multidimensional task which requires adequate resources from a technical, social, legal, and most importantly, from a human viewpoint [44, 68]. Especially the OT domain needs additional and continuous attention from the responsible stakeholders [69]. Nevertheless, 100% secure IACSs are not achievable [70]. Security and the necessary associated measures always need to be regarded with various characteristics in mind, e.g. resulting cost, effects on performance and usability, availability of human experts, and the knowledge about possible threats and attackers. Therefore, the typical risk-based approach to weigh different aspects and scale effects accordingly is followed within this dissertation. Security as the fundamental research domain is defined as follows (adapted from the IEC 62443-1-1 standard and the National Institute of Standards and Technology (NIST) glossary).

Security[1]: A condition in which risks posed by threats and vulnerabilities are reduced to an adequate level by countermeasures that enables a system to perform its mission or critical functions.

[1] https://csrc.nist.gov/glossary/term/security

The definition of security itself is based on various terms, concepts, and their relationships. Figure 3.1 further illustrates the associated contents in a visual way (based on the ISO/IEC 15408-1 standard). This dissertation uses a definition of security focusing on security **risks** (highlighted in grey within the figure below) which describe how an asset is affected by a **threat**. The potential negative consequences are characterised based on the **impact** and the **likelihood/complexity**. An **asset** as a physical hardware component with digital elements, e.g. sensors, actuators, Programmable Logic Controllers (PLCs), Human-Machine Interfaces (HMIs), or networking components, such as switches or routers, is owned by a typical **stakeholder** from the OT domain. Furthermore, assets can have technical **vulnerabilities** which are weaknesses that could be exploited by a human **attacker**. The associated stakeholders also implement certain **countermeasures** (also called mitigations or controls) related to the possible impact of a risk and reducing the overall risk score.

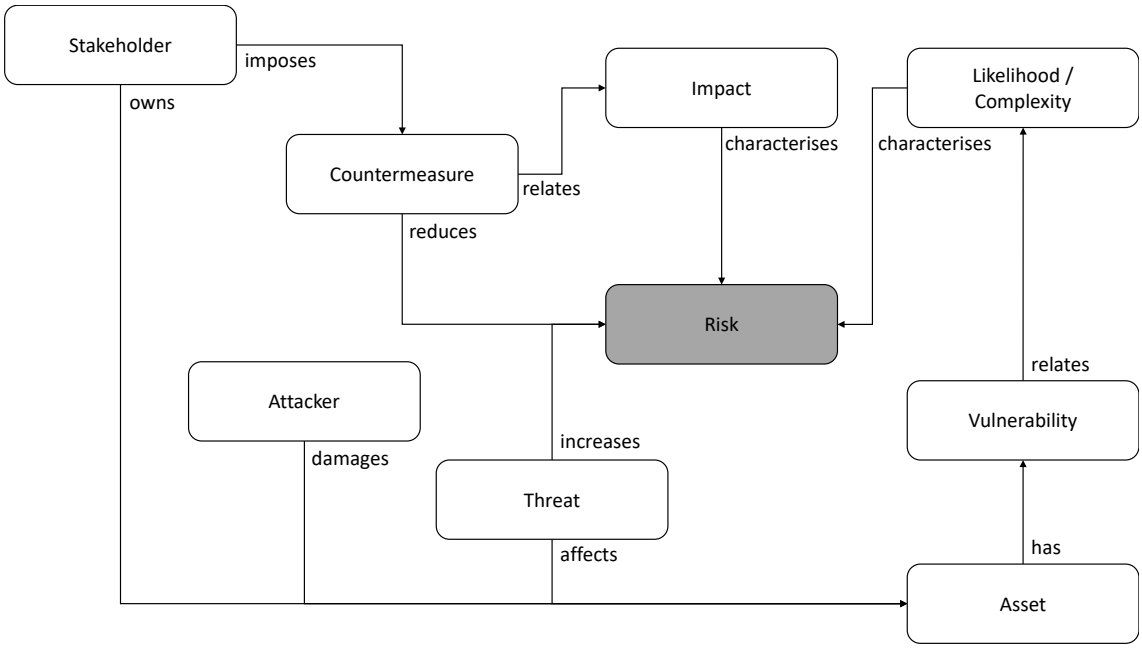

Figure 3.1: Relevant security concepts and their relationships for this dissertation (adapted from the ISO/IEC 15408-1 standard)

The OT domain consists of a plethora of stakeholders with all of them having an effect on the security of a SUC as a group of investigated assets. Figure 3.2 summarises the involved stakeholders and shows their relationships to each other. The different roles are based on the IEC 62443-1-1 and the VDI VDE 2182 standards. This dissertation focuses on the **system integrator** and **asset owner** perspectives (highlighted in grey) representing typical security-related activities for the design, commissioning, configuration, operation, and maintenance for modular manufacturing systems [71]. In addition, other roles are also involved. The **component manufacturer** is responsible for the specification and development of secure assets.

Furthermore, the **service provider** can take over dedicated tasks for an asset owner, e.g. hosting of virtual environments, configuration of specialised assets, or (remote) maintenance. **Regulatory authorities** set up legislative frameworks for the other stakeholders and a **compliance authority** is responsible for the check and certification of certain systems. Finally, **external consultants** can support or take over any of the beforehand mentioned roles within a limited scope for, e.g. the asset owner.

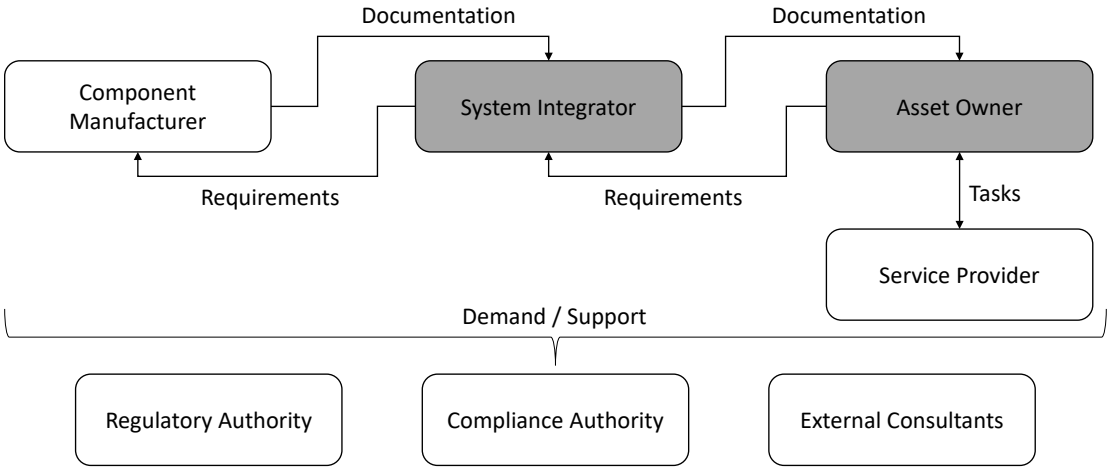

Figure 3.2: Relevant security stakeholders and their relationships for this dissertation (adapted from the IEC 62443-1-1 and the VDI VDE 2182 standards)

The Technical Report (TR) 84.00.09 by the ISA provides a RASCI (Responsible, Accountable, Supporting, Consulted & Informed) matrix for the above mentioned stakeholders showing their tasks as well as how their work is intertwined and dependant upon each other. For the scope of this dissertation focusing on the initial security risk assessment (defined within the IEC 62443-3-2) the following duties per stakeholder can be found within the ISA-TR 84.00.09:

- Asset Owner: Accountable (A)
- System Integrator: Responsible (R) & Consulted (C)

3.2. IT/OT Convergence

The definition of security, presented beforehand, is generically valid, but looking at the real world a distinction has to be made. In general, there are two domains which need to be taken into account: (1) IT and (2) OT. Historically both domains were regarded in an isolated manner due to different tasks, software implementations, used hardware components, missing interfaces and communication, and a clear distinction between stakeholders and responsibilities [72]. The current developments (also

described within the motivation in Section 1.1) show that both domains are moving towards each other and already begin to intertwine. This process is generally called IT/OT convergence and requires additional efforts regarding leadership, planning, education, communication, testing, observation, and most importantly technological improvements to keep pace with the upcoming changes [10].

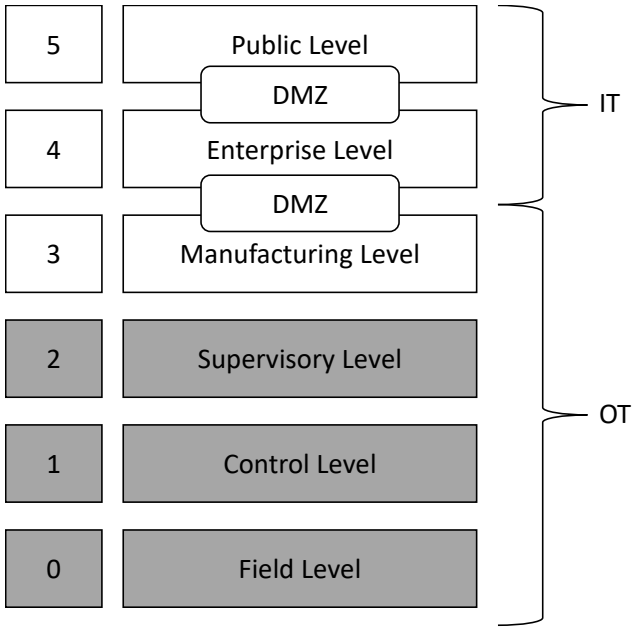

Figure 3.3: Security architecture and focus highlighted in grey as a basis for this dissertation (based on [10] and [73])

Figure 3.3 shows the current state of typical security-related network architectures based on the IEC 62264 standard, the Purdue Enterprise Reference Architecture (PERA) [10], normally referenced as the Purdue model, and the descriptions from the ISA-TR 84.00.09 [73]. The figure is an internationally agreed-upon representation of the IT/OT convergence and the associated assets. The architecture consists of six levels which can be either assigned towards the IT or the OT domain. In general, communication between the levels is only allowed towards the directly neighboured levels, e.g. the field level (0) should only be allowed to communicate with the control level (1). The top two levels (4 & 5) build up the IT domain and include Internet-based assets (typically Commercial off-the-Shelf (COTS) hardware components) within the public level, such as web or email servers, and typical business assets, e.g. standard desktop computers, domain controllers, or Enterprise Resource Planning (ERP) systems. Both levels are separated via a Demilitarized Zone (DMZ) regulating the communication interfaces and possibilities to secure the lower levels of the architecture model from attacks with, e.g. terminal servers, jump hosts, remote access, or Virtual Private Network (VPN) servers. The lower levels (0-3) are assigned to the OT domain representing manufacturing operations, containing e.g. batch,

continuous, and discrete control systems, for the production of goods [10]. They are also separated via a DMZ from the upper levels to achieve a distinction between the IT and the OT domains. Typically, the manufacturing level 3 includes assets, such as Manufacturing Execution Systems (MESs), engineering workstations, communication servers, e.g. using Open Platform Communications Unified Architecture (OPC UA), or data historians. The three levels (0-2) highlighted in grey are the focus of this dissertation. There, normally Ethernet-based communication protocols are used and the contained network architectures are typically affected by safety issues [10]. The second level normally includes Supervisory Control and Data Acquisition (SCADA) systems or local HMIs. Furthermore, the control level consists of assets, e.g. PLCs, field controllers, or industrial computers with dedicated control software. The lowest field level is directly next to the production process and typically includes sensors, actuators, frequency inverters, or Safety Instrumented Systems (SISs).

Due to this variety of asset types, communication interfaces, mixing of architecture levels, and tasks regarding IT and OT demands, the already ongoing IT/OT convergence always needs a view on the various requirements of both domains. This leads to common approaches and solutions being available for some requirements, but also to major differences between the two domains. Table 3.1 illustrates an overview of the differentiation of the security-related characteristics of the IT and OT domain.

Table 3.1: Differentiation overview and comparison of security-relevant IT and OT characteristics [20, 44, 72, 74]

Category	**IT**	**OT**
Availability	Deficiencies can often be tolerated Starts, stops, and reboots are possible	High requirements may necessitate redundant systems Outages for maintenance must be scheduled
System Lifetime	3-5 years	15-20 years and more
Performance	Non real-time High throughput Emergency interaction is less critical	Real-time Medium throughput Emergency interaction is critical
Resource Constraints	High asset complexity Enough capacity for additional applications	Low asset complexity Designed only to support the intended use case
Actuality	Very frequent updates State of the art	Patching happens rarely or not at all Legacy assets
Impact	Information and money No safety issues	Human lives, products, environment, and money Safety-critical
Approaches	Facilities accessible locally Standardised technologies Homogeneous solutions (COTS) Centralised star topology	Isolated and remote facilities Proprietary technologies Heterogeneous, individual, and embedded systems System-specific, local, and flat networks
Network Scanning	Possible, often done actively	Always cautious and if at all mostly passive
Asset Inventory	Mostly available	Rarely available and up-to-date
Risks	Manage data Delay of business operations Momentary downtime is not a major risk	Control physical world Regulatory non-compliance Loss of life, equipment, or production (safety)
Security Awareness	High	Low

The beforehand mentioned security-related requirements from Table 3.1 show a clear distinction between the two domains although the IT/OT convergence is already happening today. Therefore, enhanced approaches and concepts are required to represent the convergence of both domains [44], e.g. the OT domain still has the specific requirements but is endangered from new threats originating within the IT

domain. This also holds true for general best practices regarding security, such as engineering, configuration, management, risk assessments, threat classifications, or vulnerability analyses [74]. This dissertation focuses on the OT domain and the automation of the corresponding security risk assessments. The stated requirements and characteristics for the OT domain are taken into account for the proposed solution approach.

A special part of the IT/OT convergence are the security objectives of each domain. These general goals need to be achieved to succeed in ensuring appropriate security. Due to the differences between IT and OT characteristics, distinctive security objectives for each domain are present as well. Table 3.2 shows a comparison between the importance of the three main security objectives by ranking and ordering them [20, 44, 72, 75]:

Table 3.2: Comparison of security objectives of the IT and OT domains

Ranking	**IT**	**OT**
1.	Confidentiality	Availability
2.	Integrity	Integrity
3.	Availability	Confidentiality

The three general security objectives are confidentiality, integrity, and availability, which in combination are also called the CIA triad. They can be specified as follows by taking the definitions from the NIST glossary:

- **Confidentiality[2]:** Preserving authorised restrictions on information access and disclosure, including means for protecting personal privacy and proprietary information. Subcategories are unlinkability, untraceability, unobservability, obscurity, anonymity, or plausible deniability.
- **Integrity[3]:** Guarding against improper information modification or destruction and includes ensuring information non-repudiation and authenticity. Subcategories are accountability, authenticity, reviewability, or non-propagation.
- **Availability[4]:** Ensuring timely and reliable access to and use of information. Subcategories are dependability, reliability, or controllability.

The availability of modular manufacturing systems producing goods is of utmost importance and has the highest priority regarding the security objectives for the OT domain [74]. In addition, it is tightly linked with the integrity of data to ensure a correct, safe, and secure production process. In comparison to the IT domain, confidentiality only plays a minor role [44]. On the other hand, the security objectives

[2]https://csrc.nist.gov/glossary/term/confidentiality
[3]https://csrc.nist.gov/glossary/term/integrity
[4]https://csrc.nist.gov/glossary/term/availability

regarding IT environments have an exchanged order resulting in confidentiality with the highest ranking, followed by integrity, and at last the availability of systems.

In addition to the described security objectives, safety plays a crucial role for the OT domain as a general non-functional requirement, as a mandatory boundary condition, and as an essential function to maintain health of workers, the environment, and the used assets (following the IEC 62443-1-1 definition) [73, 76]. Therefore, this dissertation focuses on the safety-related characteristics of assets as the main objective for the automation of security risk assessments. In addition, the SEMA referential framework can be used to further narrow down the security and safety definitions utilised within this dissertation [29, 77]. Figure 3.4 displays the associated definitions based on the S–E (system & environment) and M–A (malicious & accidental) distinctions. Security covers the protection regarding malicious activities by an external threat within the environment towards the SUC and the lateral movement between systems. Whereas, safety includes the protection regarding accidental impacts from the SUC towards the environment, e.g. a human operator or nature, or towards another system, e.g. different assets or modules. Further important definitions and specifications can be found within the upcoming sections and especially within the method scoping in Section 5.2.

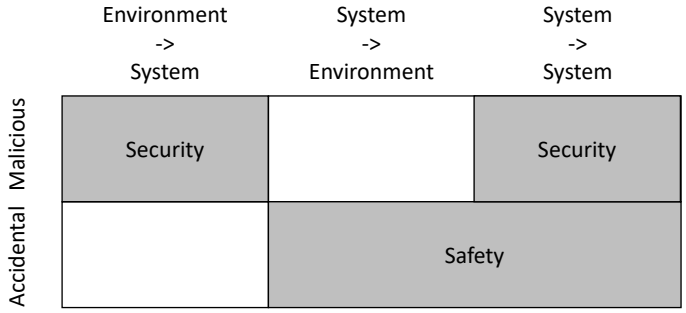

Figure 3.4: Definition of security and safety following the SEMA framework

3.3. Legal Requirements

All security-related goals, requirements, objectives, and demands are typically coming from associated legislative frameworks which specify the rules to follow for a certain domain or region. This section gives an overview of the associated legal activities surrounding this dissertation. The focus is especially set on the German region and the corresponding security-related rules coming from national and European lawmakers. The topic of security is currently in a changing condition and gets a lot of additional attention due to market pressure and a rising amount of incidents (see Section 3.5 later on). Figure 3.5 provides a high-level overview of the legislative frameworks in scope and the displayed laws and activities are sorted by their year of publication in a chronological order and their importance for this dissertation

based on a subjective evaluation. It can be shown that special areas, such as critical infrastructures or governmental agencies, are already regulated. Other areas, e.g. the production and manufacturing domain which is in scope of this dissertation, are within the regulation process at the moment. In contrast to that, the related safety domain already has a strong legal background since around 20-30 years and therefore specific demands for components and systems, e.g. to be sold and operated within the European Union (EU) via the Conformité Européenne (CE) mark.

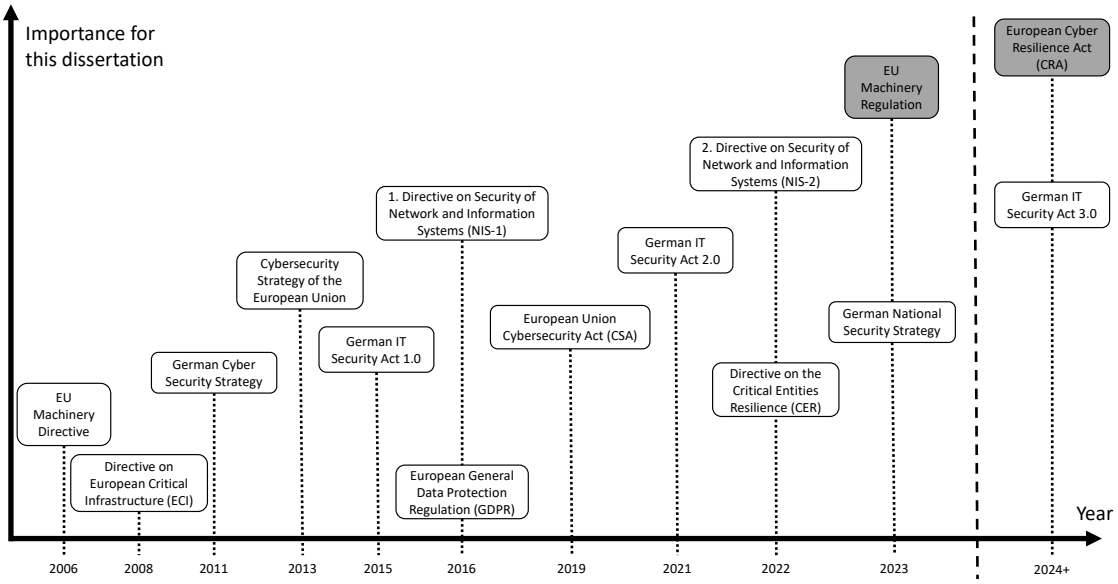

Figure 3.5: Overview of the legislation from the security and safety domains based on the publication year and the importance for this dissertation

The first directives which are worth to be mentioned date back to the mid 2000s covering mainly safety-related issues, especially with the EU Machinery Directive[5], and security-related issues for critical infrastructures[6]. This marked the beginning for the topics relevant for this dissertation. By 2011 and from then ongoing several general security strategies[7] and acts[8] in Germany and from the EU[9] were introduced to strengthen the overall security status and to put additional attention onto the topic. The year 2016 marks an important milestone, because the General Data Protection Regulation (GDPR)[10] was introduced to cover the topic of privacy in an extensive way and the first Directive on Security of Network and Information Systems

[5]https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006L0042
[6]https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32008L0114
[7]https://www.enisa.europa.eu/media/news-items/german-cyber-security-strategy-2011-1
[8]https://www.bsi.bund.de/EN/Das-BSI/Auftrag/Gesetze-und-Verordungen/IT-SiG/1-0/it_sig-1-0_node.html
[9]https://ec.europa.eu/digital-single-market/en/news/eu-cybersecurity-plan-protect-open-internet-and-online-freedom-and-opportunity-cyber-security
[10]https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

(NIS-1)[11] took effect, putting the topic of security on the agenda of all EU member states and providing a first comprehensive legislative framework. Several other acts[12] followed and in 2022 the second Directive on Security of Network and Information Systems (NIS-2)[13] highly focusing on security risk management in correspondence to the Directive on the Critical Entities Resilience (CER)[14] took effect. In 2023, the EU Machinery Regulation (highlighted in grey) was finally agreed upon and replaced the outdated directive from the year 2006. This regulation is quite important for this dissertation due to the safety-related contents which demand a connection to security for the first time in a law as well. This includes especially contents from the Annex III[15] aiming at mandatory security-related features typically included in a security risk assessment process:

- 1.1.9: Protection against corruption → "The machinery or related product shall be designed and constructed so that the connection to it of another device, via any feature of the connected device itself or via any remote device that communicates with the machinery or related product does not lead to a hazardous situation."
- 1.2.1: Safety and reliability of control systems → "Control systems shall be designed and constructed in such a way that [...] they can withstand, where appropriate to the circumstances and the risks, the intended operating stresses and intended and unintended external influences, including reasonably foreseeable malicious attempts from third parties leading to a hazardous situation."

The upcoming years starting with 2024 will provide additional acts regarding the improvement of the overall security. Especially, the foreshadowing of the European Cyber Resilience Act (CRA)[16] (highlighted in grey) already brings a lot of additional drive and attention to the topic of security also for businesses operating in the area of production and manufacturing. This holds true for various stakeholders from the OT domain, such as component manufacturers, system integrators, and asset owners. The adopted and security-improved CE mark requires certain improvements to the overall security, such as hardware and software documentation, information sharing processes, secure configurations, hardening, logging, monitoring, vulnerability analyses, and a holistic security risk assessment. These additional requirements are typically achieved via a presumption of conformity by defining and cross-referencing harmonised standards which fulfil the demands and fill the solution approach with an adequate level of details. In the case of the European Cyber Resilience Act (CRA) also the IEC 62443 standard is foreseen to be utilised as a harmonised standard.

[11] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016L1148
[12] https://www.bsi.bund.de/EN/Das-BSI/Auftrag/Gesetze-und-Verordungen/IT-SiG/2-0/it_sig-2-0_node.html
[13] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022L2555
[14] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022L2557
[15] https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32023R1230
[16] https://www.europarl.europa.eu/doceo/document/TA-9-2024-0130_EN.html

Summarising, these three findings from the legislative requirements support the contents and the scope of this dissertation:

1. Security and safety coupling is increasingly important and receives attention within the industry, among researchers, and in law making.
2. Security risk assessments are an integral part of the mandatory activities for industrial stakeholders, such as asset owners or system integrators.
3. A conformity towards the IEC 62443 standard series proves to be adequate and future-proof for the proposed solution approach of this dissertation.

3.4. Security and Safety Coupling

The coupling of security and safety is an ongoing issue due to various reasons in a plethora of industrial areas, such as aerospace, automotive, railway, energy, and especially for manufacturing and production [2, 29]. In all these applications, safety is a fundamental requirement to acquire the license to operate industrial systems as demanded by the regulative frameworks [37, 40]. As long as there are potential hazards, asset owners must reduce them to an acceptable level of residual risks [40]. This can typically be achieved by inherently safely designed systems, additional documentation, an increase of operator awareness, or by implementing safety measures, e.g. via SISs [6]. SISs supervise defined target values of industrial processes and transfer the production system into a safe state in a controlled manner if the target values exceed the tolerable range [44]. The German NAMUR recommendation NA 35 from 2019 quantifies the increased usage of SISs by a factor of three during the past ten years resulting in roughly 10% of field devices being part of a safety function. In the past, safety incidents were mainly based on accidental component failures or human errors. But today due to the current digitalisation, security-related attacks and deliberate manipulation of safety measures need to be regarded as well in this manner [44, 78]. OT personnel needs to understand the coupling of the security and safety domains as well as any impairment of one of the domains is unacceptable [72]. Furthermore, safety and security have many similarities, which can be used to align processes, combine approaches, make common methodologies more efficient, and consequently reduce the needed resources [29, 31, 78, 79]:

- Objective of preventing risks
- Establishing and fulfilling requirements
- Imposing constraints and implementing protective measures
- Enable a resilient operation of OT systems
- Common data sources, e.g. system or asset information

The coupling of security and safety approaches, e.g. for risk assessment or combined lifecycles, is a hot research topic right now. Various results can be found in the associated research publications [29, 37, 44, 77, 78, 80–86] and can be consulted for further reference. Additionally based on the regulative frameworks mentioned in Section 3.3 before, several European Norms (ENs) are harmonised under the EU Machinery Directive from 2006[17], e.g. the EN ISO 12100 as a basic machinery safety standard defining the safety risk assessment and reduction steps, the EN ISO 13849 typically used for the Performance Level (PL) determination regarding safety functions, and the EN IEC 62061 normally utilised for the Safety Integrity Level (SIL) calculation regarding safety risk scores. Furthermore, there are other internationally accepted standards, such as the IEC 61508 demanding a coupled investigation of safety and security [87], which are not harmonised under the EU Machinery Directive from 2006, but are taken if no applicable harmonised standard is available for a certain use case. The safety standards are mentioned here for completeness and to understand the bigger picture. Nevertheless, this dissertation solely focuses on the IEC 62443 standard and the corresponding definitions regarding security risk assessments with safety in mind as the most important asset characteristic.

In addition to the normative security and safety requirements, various other non-standardised documents are available in the form of TRs. These pick up the current topics and issues emerging from the coupling of security and safety to formulate the ongoing discussions, ideas, concepts, and solution proposals in a non-binding manner open for discussion within the respective communities. Several organisations, such as the IEC/TC65 plenary board, try to couple both domains at the moment [88]. Here, a short overview of the most fitting ones for this dissertation is provided: ISA-TR 84.00.09 from 2023, IEC TR 63069 from 2019, IEC TR 63074 from 2017, and ISO TR 22100-4 from 2018.

The most important and most current document is the ISA-TR 84.00.09 from 2023 which covers a coupled lifecycle for security and safety based on the ISA 61511 and the IEC 62443 standards including an exemplary process and the associated methodologies [31]. In the past, a safety-relevant Process Hazard Analysis (PHA) excluded security attacks from the intended scope of inspection regarding safety risk assessments. But due to the ongoing coupling of both domains and the increasing usage of technical solutions, such as SISs, an integration of security into the PHA is needed although this might result in less usability and convenience for the operators. Therefore, the ISA-TR 84.00.09 provides insights how to specify, implement, operate, and maintain SISs in a secure and safe manner.

Another TR in the direction of a coupled framework of security (IEC 62443) and safety (IEC 61508) for industrial-process measurement, control, and automation systems is the IEC TR 63069 from 2019. It includes general term definitions for security, safety, and their combined risks for the sake of co-engineering [44]. In addition, three guiding principles are provided, which include lifecycle recommendations

[17] https://www.pilz.com/download/open/Pos_Functional_safety_1003920-EN-05.pdf

for co-engineering of both domains, risk assessment considerations, and incident response readiness as well as incident handling [31].

A similar concept is proposed within the IEC TR 63074 from 2017 covering the safety-related security aspects and requirements, such as threats and vulnerabilities, for ICSs with the aim to maintain a safe and secure operation [88]. In addition, concrete guidances and recommendations regarding possible countermeasures implementations are given, especially for the aspect of asset integrity manipulation also anticipating the demands from the newer EU Machinery Regulation for component manufacturers, in order to couple the security and safety domains [31].

The last available TR relevant for this dissertation is the ISO TR 22100-4 which covers the general safety of machinery in relationship with the ISO 12100 standard and presents considerations for a security and safety coupling, especially for system integrators [31]. It presents similarities and differences between both domains based on certain characteristics, such as goals, conditions, approaches, inputs, outputs, or used metrics, and a five-step process (Identify, Protect, Discover, React, and Restore) to secure systems [31]. Furthermore, general advisories for the topic of remote access and maintenance are given to increase the overall security in compliance to safety requirements [31].

The last part of this section covers the coupling of security and safety as understood within this dissertation. Further information about the general scoping of the utilised method steps and associated contents can be found within Section 5.2 accordingly. Figure 3.6 shows the relationship between security and safety with the main focus highlighted in grey.

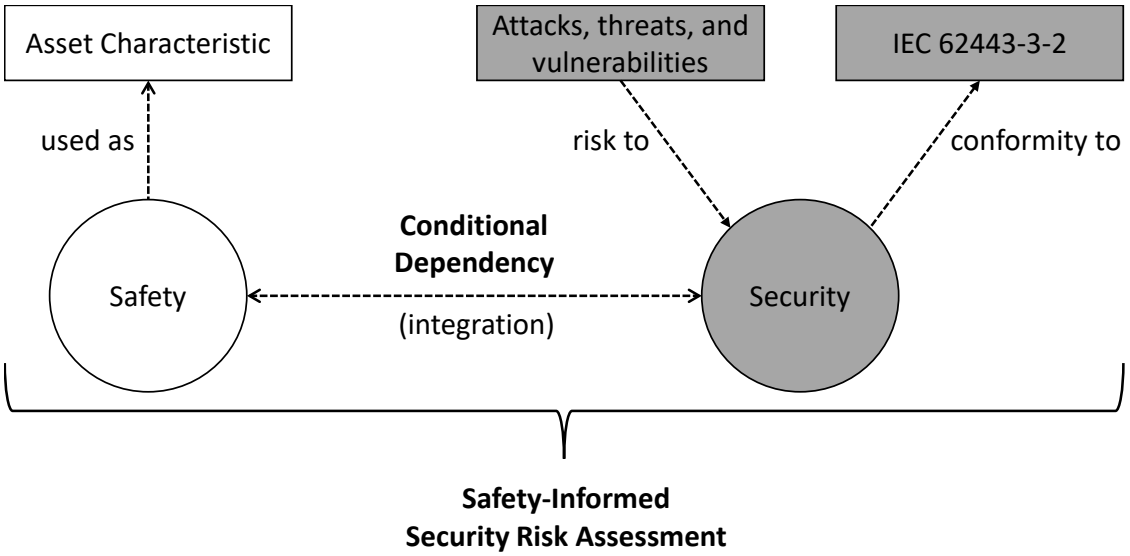

Figure 3.6: Relationship between security and safety within the context of this dissertation (adapted from [78] and based on [37, 77, 80, 84])

In general, the authors of [77] propose four types of security and safety coupling, whereas the solution approach from this dissertation can be categorised as a conditional dependency (highlighted in bold below) having reinforcing characteristics because both domains complement and strengthen each other in one focused approach.

- Mutual reinforcement: Fulfilment of safety requirements or safety measures contributes to security, or vice-versa, thereby enabling resource optimisation and cost reduction.
- **Conditional dependency:** Fulfilment of safety requirements conditions security or vice-versa.
- Antagonism: When considered jointly, safety and security requirements or measures lead to conflicting situations.
- Independency: No interaction at all.

The work by [84] analyses a variety of security and safety coupling approaches and deducts a three-tier categorisation. This dissertation can be grouped into the category of safety-informed security approaches (highlighted in bold below) by adapting safety as the most important asset characteristic for the network segmentation, asset identification, and as an objective for the security risk assessment itself. This ensures a thorough analysis of the security domain with safety in mind [37].

- Security-informed safety approaches: Approaches that extend the scope of safety engineering by adapting cybersecurity-related techniques.
- **Safety-informed security approaches:** Approaches that extend the scope of security engineering by adapting safety-related techniques.
- Combined safety and security approaches: Combined approaches for safety and cybersecurity co-engineering.

Another classification is described within [29] which distinctions between unified and integrated approaches for the coupling of security and safety. The proposed solution approach of this dissertation uses similarities of both domains and tries to align them without turning them into one single methodology which results in an integrating approach (highlighted in bold below). Safety is used as part of the security risk assessment process. It improves the overall result quality, supports a qualitative approach, and frames the scope of the complete process [29, 83].

- Unification approaches are aimed at uniting safety and security techniques into a single methodology. The result of these approaches is a single set of requirements describing the safety and security functions of the system.

- **Integration or harmonising approaches** are aimed at investigating the similarities and differences of safety and security techniques in order to bring them into alignment. These approaches produce safety and security requirements separately using standard concepts and methodologies, and then show how they interact in order to identify conflicts.

3.5. External and Environmental Aspects

In addition to the created field of tension due to the IT/OT convergence, the present legislative requirements, and the coupling of security and safety, this section collects facts about the overall security-related environment this dissertation is placed within. The displayed status is based on the provided overview of incidents, surveys, and general challenges.

The ongoing IT/OT convergence results in various new attack vectors causing a plethora of incidents, especially within the manufacturing and production domain, resulting in an estimated global loss of about 6 trillion € in 2021 for the global economy [89]. Bitkom, a digital association from Germany, states that the German economy had a damage of around 203 billion € in 2022[18]. Already in 2020 the World Economic Forum ranked cyber attacks and critical information infrastructure breakdown as the most dangerous global technological risk and the second most concerning risk for doing business globally over the next 10 years [90]. In addition, the IBM Security Threat Index from 2022 shows that the manufacturing industry ranks on the first place for security-related attacks with 23.2% [91]. Furthermore, Kaspersky states that attacks on the industrial sector hit a record high in the second quarter of 2023 with 26.8% of all OT-related computers being affected by malicious objects[19]. This development is mainly due to interferences with assets, stolen passwords, data breaches, espionage, or ransomware deployed by around 83% of external threat actors according to the Verizon Data Breach Investigations Report (DBIR) from 2023[20]. The results are typically disruption of operation, failure, theft or damage to information systems or operating processes, loss of sales, or loss of competitive advantage [72]. Table 3.3 provides an overview of security-related incidents relevant for this dissertation to give an impression of the chronological developments, the affected domains, and the associated components (without aiming for completeness). The collection of incidents originates from the OT, safety, and critical infrastructure domains.

[18] https://www.bitkom.org/Presse/Presseinformation/Wirtschaftsschutz-2022

[19] https://www.kaspersky.com/about/press-releases/2023_attacks-on-industrial-sector-hit-record-in-second-quarter-of-2023

[20] https://www.verizon.com/business/resources/reports/dbir/2023/summary-of-findings

[21] https://www.pilz.com/en-US/company/news/articles/215551

[22] https://www.techtarget.com/whatis/feature/SolarWinds-hack-explained-Everything-you-need-to-know

[23] https://www.cisa.gov/news-events/news/attack-colonial-pipeline-what-weve-learned-what-weve-done-over-past-two-years

Table 3.3: Overview of security-related incidents framing the security environment of this dissertation

Year	Description	Domain	Reference
1982	Trans-Siberian pipeline	Critical infrastructure	[64]
2010	Stuxnet	Critical infrastructure	[38]
2014	German steal manufacturing plant	OT	[92]
2015	BlackEnergy3	Critical infrastructure	[93]
2016	CrashOverride	OT	[93]
2017	TRITON/TRISIS/HatMan	Safety	[94, 95]
2019	German Manufacturer Pilz	Safety	Link[21]
2020	SolarWinds	Critical infrastructure	Link[22]
2021	Colonial Pipeline	Critical infrastructure	Link[23]
2023	Toyota shuts down 14 Factories	OT	Link[24]
2023	PLC exploitation	OT	Link[25]
2024	ThyssenKrupp Automotive	OT	Link[26]

To further emphasise the importance of automated security risk assessments, the following list of surveys serves as an insight into the OT domain and the typical conditions at companies, especially SMEs. This lays the foundation for the contents and achievements of this dissertation. The statements from the collected and displayed surveys are based on different data sizes and were performed by different organisations. Nevertheless, the provided statements within the following list can be used to further understand the security-related environment in scope:

- **Preconditions**
  - 100% of organisations investigated had routable network connections into their OT environments [96]
  - 66% of attacks towards OT networks involved adversaries directly accessing them from the Internet [97]
  - 71% of assets have outdated/unsupported Operating Systems (OSs) [98]
  - 89% of companies have business critical systems within their OT environments [76]
  - 44.8% of OT systems allocate less than 1% of the security budget [99]
  - 78% of typical OT organisations were already victim of a security incident [100]

[24] https://www.reuters.com/business/autos-transportation/what-happened-shut-down-toyotas-production-japan-2023-08-30/
[25] https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-335a
[26] https://www.bleepingcomputer.com/news/security/steel-giant-thyssenkrupp-confirms-cyberattack-on-automotive-division

- **Security**
  - 71% of organisations assessed had poor a security perimeter [101]
  - 83% of companies know one of the common security standards, e.g. IEC 62443, with just 41% applying it [102]
  - Only 20% of industrial organisations use OT-related security policies [103]
  - Security compliance takes up more than 30% of the investment resources in smaller organisations, e.g. SMEs [104]
  - Only 44% of companies perform security risk assessments on a regular basis [76]
- **Safety**
  - Around 67% of safety-relevant incidents included a correspondence to security [76]
  - Highly impactful public OT exploits affect every level of industrial environments [105]
  - The coupling of security and safety is seen as an obstacle and issue [100]
- **Impacts**
  - In 2017, security incident costs per organisation were between 68.880 € for SMEs and 1.5 Million € for corporation groups [106]
  - In 2021, the average cost of a data breach within the industrial sector was around 4 Million € [107]
  - Top five countries affected by ransomware and extortion attacks are the US, the UK, Germany, Canada, and France [108]
  - 45,5% of companies want to increase visibility into their OT-related control systems [109]

In addition to the security-related incidents and surveys, there is a plethora of challenges present within the OT domain regarding the implementation of security risk assessments. Table 3.4 illustrates an overview of the security-related challenges in scope for the automation of security risk assessments and can be seen as an addition towards the challenges already mentioned within Section 1.3. The challenges are listed and categorised based on their focus [92] (People / Process / Technology) and their relevance (Low / Medium / High) for this dissertation. The provided summary underlines the need for automated security risk assessments and plays a crucial role in the further understanding and motivation of the proposed solution approach.

Table 3.4: Overview of security-related challenges framing the overall security environment of this dissertation

Challenges	Focus [92]	Relevance	Reference
Missing coverage of all security risk assessment stages	Process	High	[64, 92, 110]
Supporting risk assessment methods with elaborate software tools	Technology	High	[55, 64]
Lack of security and safety coupling	Process	High	[30, 83]
Lack of generally applicable and industry-compatible standards	Process	High	[111]
Support of security for dynamically reconfigurable automation systems	Technology	High	[83]
Integration of the digital twin in security management	Technology	High	[55, 112, 113]
Continuous compliance monitoring with a high resource consumption	Process	High	[111]
OT security shall be ensured throughout the entire lifecycle	Process	High	[44]
Provisioning and modelling of OT security knowledge	Technology	High	[44, 65, 114]
Opaque security definitions and limited targets of evaluation	Process	Medium	[115]
Rise of high-impact vulnerabilities and complexity of underlying systems	Technology	Medium	[30, 55, 114, 116]
Unknown probabilities of security-related events, e.g. human attackers	Process	Medium	[2]
Assigning asset values and estimating possible impacts	Process	Medium	[117]
SMEs need special support for OT security analyses	Technology	Medium	[44]
Common understanding of the IT/OT convergence	People	Medium	[44]
Guarantee of transparent and credible results with a steady quality	Process	Medium	[118]
Integration of the human factor and the corresponding impacts	People	Low	[55, 92, 110]
Scarcity of historical and reference data sets for simulations and learning	Technology	Low	[55, 64, 65, 110]
Fear of increased system complexity due to security measures	People	Low	[111]
Unclear contribution of security investments to value creation	People	Low	[111]
Raising security awareness due to lack of understanding and knowledge	People	Low	[119]

3.6. Security Risk Assessment Processes

Based on the IT/OT convergence, the legislative frameworks, the security and safety coupling, and the prevailing security environment, security is a generally mandatory requirement for any OT environment today originating from various sources, such as the different industrial stakeholders, standards and guidelines, governmental agencies, literature and research, public databases, and Computer Emergency Response Team (CERT) notifications [44, 120]. This is also supported by the motivation (in Section 1.1) and the proposed solution approach (in Section 1.4) of this dissertation.

Nevertheless, there is never a 100% guarantee to build a secure system [70] which meets the demanded security levels and the desired security objectives [20]. Therefore, security needs to be part of the general strategy of every organisation [74]. Security-related standards, best practices, and regulations are proposed on a global scale supporting organisations to achieve their goals [20]. These proposals try to provide frameworks for the responsible personnel to perform assessments, evaluations, audits, and hardening with their systems to be secure against the majority of attacks or at least to become unrewarding for possible adversaries [121]. In general, this helps organisations to describe the current and target states regarding security, identify and prioritise opportunities for improvement, evaluate the corresponding processes, and communicate the findings towards all relevant stakeholders [121].

The common recommendation for IT and OT environments alike is to perform security-related activities in a continuous manner, e.g. following the general Plan, Do, Check and Act (PDCA) cycle or the Observe, Orient, Decide, and Act (OODA)

approach [30]. Regarding security, this is typically done via organisational and technical processes to fulfil GRC requirements by implementing an Information Security Management System (ISMS) based on the three pillars of people (e.g. awareness, training, or incident response exercises), processes (e.g. risk management, defense in depth, or incident response plans), and technology (e.g. firewalls, Intrusion Detection Systems (IDSs), Intrusion Prevention Systems (IPSs), or Security Information and Event Management (SIEM)) [44, 122]. The ISMS concept is generally defined within the ISO/IEC 27000 standard series and comes originally from the IT security domain. Other domains, such as Trusted Information Security Assessment Exchange (TISAX) for automotive products which define Cyber Security Management Systems (CSMSs), took over the basic concept and adapted it with specific characteristics for the respective application. A usable concept regarding ISMSs for the OT domain is proposed within the IEC 62443-2-1 standard under the name of an IACS Security Program (SP) [123].

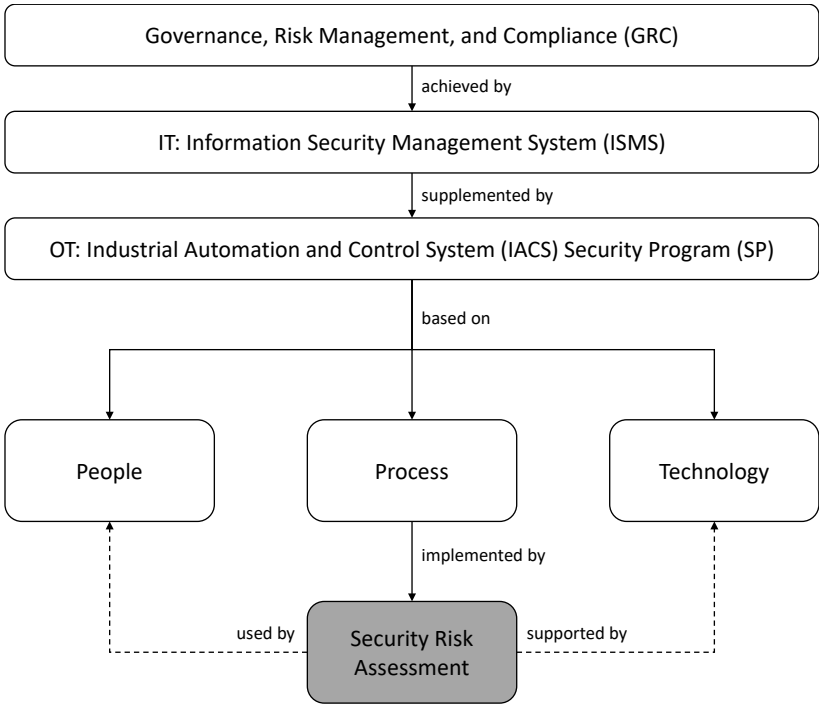

Figure 3.7: Security risk assessments as a focus of this dissertation highlighted in grey in alignment with the general GRC requirements

In general, risks are inherent in all aspects of every organisation independent of the domain or the size [2, 44]. Therefore, security risk assessments are an integral and central part of the ISMS concept including risk identification, risk analysis, and risk evaluation as generally specified within the ISO 31000 standard or the NIST Cybersecurity Framework (CSF) approach [44]. Further references for security risk assessment processes are the ISO 31010, NIST 800-82, NIST 800-30, or NIST 800-39.

Other less popular standards can be found within the extended literature [39, 64, 124–128]. The dedicated security risk assessment process for the OT domain is defined within the IEC 62443-3-2 standard. Figure 3.7 illustrates the scope highlighted in grey in alignment towards the specified concepts.

The history of generally using risks as a qualitative or quantitative measure for real-world problems is quite old, e.g. shown in [129] from 1954 or in [130] from 1981. Following the long trail of publications regarding this topic and the associated definition of risks, e.g. the works of [29, 44, 62, 65, 131], a current one can be found within the NIST glossary. Instead of a fixed definition, this dissertation uses the following function to describe a security risk based on threats (T), technical vulnerabilities (V), attacker skills and techniques (A), and the safety characteristic of assets (S). The details of the utilised security risk specification can be found within Section 6 to Section 9:

Risk[27] = f (T, V, A, S)

Based on the functional specification of a security risk for this dissertation, the following definition of a security risk assessment is used throughout the document. It is based on the understanding from the ISO 31000 standard reflecting the process in a three-tier way consisting of risk identification, risk analysis, and risk evaluation. Risk identification is the process of recognising and describing the conditions under which adversaries can cause an impact [114]. The risk analysis has the goal to review the identified risks and to provide value-based estimates for the impact and the likelihood/complexity of a specific risk [114]. Finally during the risk evaluation, it is determined how the identified and analysed risks are going to be treated [114]. In addition, the definition is enhanced with the contents from the NIST glossary:

Security Risk Assessment[28]: Overall process of risk identification, risk analysis, and risk evaluation to comprehend the nature of risks and to determine the level of risks based on threats, vulnerabilities, attackers, and safety regarding the operation of modular IACSs.

The described focus on security risk assessment processes is backed by various factors, such as the already wide distribution of usage [76, 109], the high priority to ensure availability of industrial systems [109], the easy implementation in a qualitative manner [54], the characteristic of a key starting point for any security-related activity [132–134], and the high need for additional technical assistance for assessments via tools based on security-relevant information [99]. In addition, security risk assessments are necessary for well-informed management decisions due to low budgets and need for adequate prioritisation of resources [2]. Organisations which do not perform these assessments on a regular basis may experience severe consequences for their systems during security incidents resulting in, e.g. loss of reputation, legal

[27] https://csrc.nist.gov/glossary/term/risk
[28] https://csrc.nist.gov/glossary/term/risk_assessment

issues, or even a direct financial impact [48]. Furthermore, an appropriate risk management favours the understanding of a common business strategy and facilitates communication across the different stakeholders and the responsible personnel.

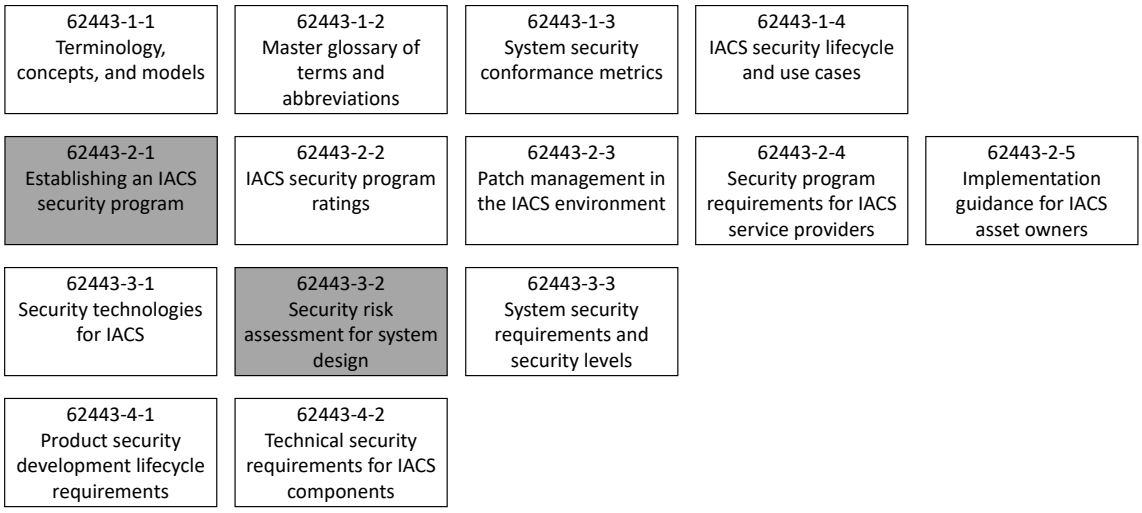

Figure 3.8: Overview of the IEC 62443 suite of standards for the OT domain including the focus of this dissertation highlighted in grey

Today, the most comprehensive security standard for the OT domain is the IEC 62443 suite published by the ISA. Therefore, the focus of this dissertation is placed onto the IEC 62443 suite and the associated security risk assessment definitions from the IEC 62443-3-2. The given contents there are influenced by various other standards, guidelines, and numerous globally distributed organisations, such as the ISO/IEC 27000 standards series, the German VDI/VDE 2182 guideline, several advisories from the German Bundesamt für Sicherheit in der Informationstechnik (BSI), the Critical Infrastructure Protection (CIP) approaches by the North American Electric Reliability Corporation (NERC), the special publications from the NIST, and input from, e.g. the American Department of Homeland Security (DHS). The overall standardisation landscape at the moment is very broad with many different stakeholders and various proposals, but the IEC 62443 is the agreed on de-facto standard for the OT domain. Figure 3.8 summarises the various parts of the IEC 62443 standard suite[29] including their name and identifier. The focus of this dissertation is highlighted in grey and is set on the IEC 62443-2-1 defining an IACS SP including the general motivation and necessity for risk management and the associated security risk assessment process definition from the IEC 62443-3-2 part. By doing so, a support for asset owners and system integrators is achieved.

The IEC 62443-3-2 standard specifies the security risk assessment procedure regarding the system design for IACSs. The whole procedure is based on the

[29]https://isagca.org/isa-iec-62443-standards

application of zones (grouping of logical or physical assets sharing common security requirements) and conduits (logical grouping of communication channels between zones sharing common security requirements) as a fundamental concept from the IEC 62443-1-1. In general, the standard describes how to define the investigated system as the SUC, to partition it into zones and conduits, to establish a Security Level (SL) Target, to assess the corresponding risks, and to document the results. The whole process consists of seven main ZCRs (1-7), each representing a high-level step needed for a holistic security risk assessment and being divided by a varying amount of ZCRs. Figure 3.9 illustrates the overview of all ZCRs from the IEC 62443-3-2 forming the complete security risk assessment process. The focus of this dissertation is highlighted in grey. For further explanation about the scope of contents, refer to Section 5.2. The ZCR 2 generally describes initial security risk assessments and is taken as a basis. Furthermore, it is complemented with the five most resource-intensive ZCRs (3.1, 5.1, 5.2, 5.6 & 7.1) based on the manual evaluation from Section 5.1.2.

ZCR 1 Identify the SUC	ZCR 2 Initial cyber security risk assessment	ZCR 3 Partition the SUC into zones and conduits	ZCR 4 Risk comparison	ZCR 5 Perform a detailed cyber security risk assessment	ZCR 6 Document cyber security requirements, assumptions, and constraints	ZCR 7 Asset owner approval
Identify the SUC perimeter and access points	Perform initial cyber security risk assessment	Establish zones and conduits	Compare initial risk to tolerable risk	Identify threats	Cyber security requirements specification	Attain asset owner approval
		Separate business and IACS assets		Identify vulnerabilities	SUC description	
		Separate safety related assets		Determine consequences and impact	Zone and conduit drawing	
		Separate temporarily connected devices		Determine unmitigated likelihood	Zone and conduit characteristics	
		Separate wireless devices		Determine unmitigated cyber security risk	Operating environment assumptions	
		Separate devices connected via external networks		Determine SL-T	Threat environment	
				Compare unmitigated risk with tolerable risk	Organisational security policies	
				Identify and evaluate existing countermeasures	Tolerable risk	
				Reevaluate likelihood and impact	Regulatory requirements	
				Determine residual risk		
				Compare residual risk with tolerable risk		
				Identify additional cyber security countermeasures		
				Document and communicate results		

Figure 3.9: Overview of all ZCRs specified within the IEC 62443-3-2 standard and the focus of this dissertation highlighted in grey

The IEC 62443-1-1 standard describes possible threats or attacks in a five stage scaling system in which each stage is stated as a SL: **SL 0** offers no protection at all, **SL 1** delivers protection against casual or coincidental violation, **SL 2** provides protection against intentional violation using simple means, **SL 3** gives protection against intentional violation using sophisticated means, and finally **SL 4** is described by protection against intentional violation using sophisticated means with extended resources. The SLs are designed in the way of using the attacker's motivation and

resources. Generally there is a differentiation between three types of SLs which are used throughout the security risk assessment process [135]:

- **Capability SL (SL-C):** Security level that components can provide
- **Target SL (SL-T):** Desired level of security for a particular system
- **Achieved SL (SL-A):** Actual level of security for a particular system

Furthermore, the IEC 62443 standard defines seven different Foundational Requirements (FRs) providing an abstracted view on the general security objectives for the OT domain. Depending on the use case, these FRs are further detailed by assigned System Requirements (SRs) for IACSs, such as a modular manufacturing system, or by assigned Component Requirements (CRs) for individual assets, such as PLCs, firewalls, switches, or routers. These different types of requirements are used to describe the needed characteristics for a component (CR) or for a system (SR) on a specific level in contrast to the general description of the FRs. This enables component manufacturers and system integrators to use the IEC 62443 standard to identify the requirements to secure their products accordingly. An exemplary utilisation of FRs, SRs, and CRs can be found within the Annex B in Section 15. The following list shows a summary of the seven FRs [135]:

- **FR 1:** Identification and Authentication Control - Identify and authenticate all users (humans, software processes, and devices) before allowing them to access the control system.
- **FR 2:** Use Control - Enforce the assigned privileges of an authenticated user (human, software process, or device) to perform the actions on the IACS and monitor the use of these privileges.
- **FR 3:** System Integrity - Ensure the integrity of the IACS to prevent unauthorised data or information manipulation.
- **FR 4:** Data Confidentiality - Ensure the confidentiality of information on communication channels and in data repositories to prevent unauthorised disclosure.
- **FR 5:** Restricted Data Flow - Segment the control system via zones and conduits to limit the unnecessary flow of data and lateral movement.
- **FR 6:** Timely Response To Events - Respond to security violations by notifying the proper authority, reporting needed evidence of the violation, and taking timely corrective action when incidents are discovered.
- **FR 7:** Resource Availability - Ensure the availability of the control system against the degradation or denial of essential services.

This part should serve as an introduction to the IEC 62443 concepts and summarises the main definitions which are used throughout this dissertation. For further reference, please consult the original standard documents.

4. State of the Art

The upcoming chapter contains a comprehensive view on the state of the art surrounding the contents of this dissertation. First, the next Section 4.1 presents an overview of the available tools and approaches for security engineering in general. In addition, a categorisation of tools is developed and the metric called Level of Autonomy (LOA) is introduced to measure the degree of automation and the possible support for human operators regarding technical solutions. Afterwards, Section 4.2 summarises the present formalisms which can be used for the general modelling of security and how the desired expert system is achieved as one goal of this dissertation. In Section 4.3 the related work regarding the automation of security risk assessments is presented. The overall evaluation is based on the identification of certain characteristics which are used to compare the specific works with each other. Finally, the related work is summarised and four main deficits are derived which build the foundation for this dissertation.

4.1. Landscape of Tools and Approaches

The main part of this section is the market analysis of available tools and approaches for OT-related security risk assessments which belong to the state of the art forming the environment of this dissertation. This section builds upon previous work by the author of this dissertation and extends the published results from [53]. The surrounding literature contains various surveys [2, 136, 137] and although there are multiple hundreds of tools and approaches available, research and development regarding security risk assessments are still ongoing and a recent hot topic [2, 44]. Therefore, a categorisation is described here to illustrate the different types of tools and to give an overview of the associated tooling landscape [53]: (1) Documentation, (2) checklists, (3) SIEM, (4) passive monitoring, and (5) active scanners. The idea is to propose a categorisation and to enhance these categories with the corresponding automation degree and the covered security risk assessment steps (ZCRs) from the IEC 62443-3-2 standard. The results are shown here in the following paragraphs. The evaluation of the automation degree is based on the taxonomy of LOA consisting of six levels [138]. It is based on the two key dimensions with regard to the scope of the automated tasks and the role of the human operator. The taxonomy can also be interpreted as an indicator on how much creativity (equalling a certain unpredictability) is required and how easy tasks can be formalised and repeated to be automated. The following list defines the six levels (0-5) in an abstract way [138]:

- **LOA 0:** No autonomy, humans are in full control and responsibility of the system without any assistance.
- **LOA 1:** Assistance with or control of subtasks, humans are always responsible, specifying set points.
- **LOA 2:** Occasional autonomy in certain situations, humans are always responsible, specifying intents.
- **LOA 3:** Limited autonomy in certain situations with alerting of issues, humans confirm act as a fallback.
- **LOA 4:** System in full control in certain selected situations, humans must still supervise.
- **LOA 5:** Autonomous operation in all situations, humans may be completely absent.

The presented tool examples in the following paragraphs are presented without any personal preference, supportive funding, or certain order [53]. Due to the proprietary and cost-intensive characteristics of the commercial tools, the following categorisation is based on the available public material, e.g. datasheets or trial versions, and the gained experience from former research and customer-related projects.

4.1.1. Categorisation of Commercial Tools

The **documentation** of general workplace-related risks and the creation of the corresponding high-level reports are typical tasks for nearly every responsible in all domains. With regard to security risk assessments this implies the documentation of the identification, analysis, and evaluation of security risks in a comprehensible manner to be compliant with the requirements from the applicable standards, e.g. the IEC 62443. Nowadays, the typical tools used for this task are the classical office software suites for writing texts, creating spreadsheets, and drawing figures. Various templates for these tasks are also freely and publicly available on the Internet. The given tools are not able to assist the human operator with certain tasks resulting in no autonomy at all [53]. This area of tasks is still addressed with informal information artefacts and tools [114]. In addition, to fill the tools with security-related information a huge amount of security expert knowledge is required [114]. The resulting ZCR coverage and LOA value for the tool category of documentation can be found below:

- **ZCR Coverage:** 5.13, 6.1-6.9, and 7.1
- **LOA:** 0

The next category of tools can be understood as an entry point to the topic of security for companies of every size and state of security. They contain standardised

checklists for self-assessments to achieve conformity in the form of a questionnaire mostly based on international standards which have to be answered by the user to give an estimation about the security level [42, 46, 55, 114, 139, 140]. The most famous ones for the OT domain are the Cyber Security Evaluation Tool (CSET)[1] and the CSF[2] tool-supported approach. Light and Right Security ICS (LARS ICS)[3] and the former GSTOOL[4] work in a similar way but lack current updates and are still in a reworking phase by German authorities. The Security Engineering Tool (SET) from admeritia[5] is a proprietary tool for method-agnostic security engineering. Another example is the Microsoft ThreatModeler[6] which was originally designed for Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privileges (STRIDE)-based Risk Management (RM) inside IT environments and the implemented concepts could theoretically be adapted towards the OT domain. The MITRE Critical Infrastructure Cyberspace Analysis Tool (CICAT)[7] can be used for adversary assessments within the domain of critical infrastructure. The safety domain offers additional various other tools, such as SISTEMA[8], WEKA Manager CE[9], safeexpert[10], TIA Selection Tool[11], or mCom One[12], to enable operators with checklists for self-assessments. In addition, the Control System Cyber Security Self-Assessment Tool (CS²SAT), the SCADA Security Assessment Tool (SSAT), and the Cyber Resilience Review Self-Assessment Package (CRR) are relevant for the OT domain [42]. Other commercial tools are, e.g. Verinice[13], Citicus ICS[14], Blade RiskManager[15], IriusRisk[16], SecurITree[17], or the Sandia National Laboratories (SNL) Risk Assessment Methodology (RAM) [141]. The presented tools can support the human operator in performing special subtasks, such as automatic questionnaire creation, documentation of the security state, and result presentation towards other

[1] www.ics-cert.us-cert.gov/assessments
[2] www.nist.gov/cyberframework
[3] https://www.bsi.bund.de/EN/Themen/Unternehmen-und-Organisationen/Informationen-und-Empfehlungen/Empfehlungen-nach-Angriffszielen/Industrielle-Steuerungs-und-Automatisierungssysteme/Tools/LarsICS/LarsICS_node.html
[4] https://www.bsi.bund.de/EN/Themen/Unternehmen-und-Organisationen/Standards-und-Zertifizierung/IT-Grundschutz/it-grundschutz_node.html
[5] https://www.admeritia.de/en/services_solutions.html
[6] www.microsoft.com/en-us/securityengineering/sdl/threatmodeling
[7] www.mitre.org/news-insights/publication/critical-infrastructure-cyberspace-analysis-tool-cicat-capability
[8] www.dguv.de/ifa/praxishilfen/practical-solutions-machine-safety/software-sistema/index.jsp
[9] www.weka-manager-ce.de/english-version
[10] www.ibf-solutions.com/en
[11] www.siemens.com/global/en/products/automation/topic-areas/safety-integrated/factory-automation/support/tia-safety-evaluation-tool.html
[12] www.tuvsud.com/en/industries/manufacturing/seguridad-de-maquinas/mcom-one—machinery-safety-compliance
[13] www.verinice.com
[14] www.citicus.com/ics-risk
[15] www.kdmanalytics.com/cybersecurity-products/blade-riskmanager
[16] www.iriusrisk.com
[17] www.amenaza.com

responsible stakeholders. The resulting ZCR coverage and LOA value for the tool category of checklists can be found below:

- **ZCR Coverage:** 2.1, 6.9, and 7.1
- **LOA: 1**

SIEM tools are generally designed to aid network administrators in computer security, intrusion detection, and incident prevention. This includes capabilities, such as collecting, analysing, presenting network- and security-related information, the integration of log files, or triggering warnings about findings. There are a lot of commercial products in this domain, but also open source tools are available, e.g. Open Source Security Information Management (OSSIM)[18], Enterprise Log Search and Archive (ELSA)[19], Collective Intelligence Framework (CIF)[20], or Sguil (as a collection of Snort, Suricata, SACNP, and Wireshark)[21]. By using SIEM tools, the human user is supported by automated log analyses and alerts within the specified scope representing the operator intents [38]. The resulting ZCR coverage and LOA value for the SIEM tools can be found below:

- **ZCR Coverage:** 1.1
- **LOA:** 2

OT monitoring tools passively supervise the whole network and are well-suited for the requirements in industrial environments because typical assets there mostly react very sensitively to disturbances, such as active scanning or penetration testing, or changes to their normal communication patterns [136]. Every direct interaction may disconnect assets or disable certain communication paths resulting in a system shutdown and consequently a loss of availability and productivity. These tools analyse various data sources, such as network traffic, asset information, or logs, and are able to detect anomalies within the regular patterns, such as behaviour or communication [136]. Newer solutions are also able to match available vulnerability information with the detected assets and to create threat scenarios up to a certain degree. Large vendors in this category are, e.g. Dragos[22], Forescout[23], or Nozomi[24] [136]. Other smaller solutions are available from, e.g. Claroty, Rhebo, Sentryo, Greenbone, Acht:Werk, Fortinet, or Microsoft [136]. In general, there is a plethora of proprietary tools available from commercial vendors[25]. The huge variety of analysis capabilities of passive OT monitoring tools are the most advanced ones for

[18] www.alienvault.com/products/ossim
[19] www.github.com/mcholste/elsa
[20] www.csirtgadgets.com/collective-intelligence-framework
[21] bammv.github.io/sguil/index
[22] www.dragos.com/platform
[23] www.forescout.com/products/rem
[24] www.nozominetworks.com/products/overview
[25] www.gartner.com/reviews/market/operational-technology-security

the industrial domain so far resulting in a sophisticated support for human operators but with a limited autonomy for certain tasks so far, such as network scanning and alerting. Nevertheless, the human operator always needs to be there to act as a fallback and cannot be completely absent. The resulting ZCR coverage and LOA value for the passive OT monitoring tools can be found below:

- **ZCR Coverage:** 5.2, 5.13 & 7.1
- **LOA:** 3

The last category of tools covers the active **scanners** for network analysis which are typically more widely spread within IT systems due to possible disturbances of the OT networks and assets [44]. Nevertheless, active scanning under predefined circumstances and controlled framework conditions can be beneficial to get an accurate and transparent overview, e.g. for penetration testing. A typical open source example of such tools is the network scanner NMAP[26]. Especially interesting for the industrial domain are active vulnerability scanners which can discover technical vulnerabilities of assets, e.g. Greenbone[27], openVAS[28], or Nessus[29]. In addition, specialised tools for the task of penetration testing are also available, e.g. the open source Metasploit[30] framework. These tools are typically used by domain experts after a high-level security risk assessment and the necessary scoping to dig deeper into certain parts of a system. Therefore, this category requires an extensive domain-specific know-how for the safe usage with OT systems and can only assist with certain subtasks while the responsibility is always lying at the human operator. The resulting ZCR coverage and LOA value for the tool category of active scanners can be found below:

- **ZCR Coverage:** 1.1 & 5.2
- **LOA:** 1

4.1.2. Research and Development

In addition to the vast market of proprietary, commercial, and open source tools, there are various other security risk assessments approaches available [2]. These are from the research domain, can be integrated from other domains into the OT domain, or are regarded as legacy at the time of writing this dissertation and mentioned here for completeness only. Typically these approaches try to address drawbacks from the previously mentioned tools due to advanced requirements for the OT domain or missing functionality [2]. The most important approaches are going to be described

[26] www.nmap.org
[27] www.community.greenbone.net
[28] www.openvas.org/index-de.html
[29] www.tenable.com/products/nessus
[30] www.metasploit.com

here. Other minor important security risk assessment approaches can be found within the extended literature [2, 48, 55, 64, 65, 73, 114].

The current landscape within the research domain offers a plethora of security risk assessment approaches with varying characteristics, base concepts, target groups, implementation degrees, documentation, and application domains [137]. This ranges from general approaches [142] which are based on typical IT tools, such as Excel, MatLab [64] and GRC management[31], up to domain-specific approaches [142], such as smart grid protection (SPARKS methodology [143] or the SEGRID research project[32]), automotive (Modular risk assessments [144], ThreatGet[33], or the MoRA Security Analyst[34] from the IUNO research project [145]), critical infrastructure (MoSaIK[35] and SICIA research projects[36]), or safety (AttackTree[37] or automated safety-relevant communication engineering [82]). The research into the direction of security is often based on detailed system modelling and the quantitative and probabilistic modelling of attacker behaviour, such as within ADVISE [55, 140], Möbius [64], CyberSAGE[38], or the Cyber Security Modelling Language (CySeMoL) [55, 140, 146]. More security-detailed approaches are also available based on the evaluation of incidents, e.g. by the Cyber Security Incident Risk Analysis (CSIRA) [147], or based on the assessment of technical vulnerabilities, e.g by the Framework for the Analysis of Security in CPS (FAST-CPS) [140].

Due to long history of security-related research and the increasing attention due to current developments, there is already a huge variety of legacy approaches by today. The STRIDE [55, 142, 145, 148] methodology is widely distributed in IT environments, but is not applicable to OT systems due to the lack of system-level assessments and the focus on very specific software libraries [114]. Damage, Reproducibility, Exploitability, Affected Users, and Discoverability (DREAD) [55, 142] and CORAS [48, 64, 114, 142, 145, 149, 150] are used for assessing security risks regarding IT systems based on the determination of likelihood and impact of a possible attack. Both are no longer actively maintained by their creators. Various other minor important approaches can be found within the associated surveys [48, 51, 55, 64, 65, 142, 149–151], e.g. OCTAVE, CRAMM, COBRA, ISRAM, TARA, FAIR, EBIOS, HTRA, CARVER, MSHARPP, SPRINT, BPIRM, RAIM, MAGERIT, MEHARI, or Trike. Nevertheless, all of them are either outdated or do not match the scope of this dissertation. They are mentioned here for the sake of completeness.

[31] www.crisam.net
[32] www.segrid.eu
[33] www.threatget.com
[34] www.itemis.com/de/yakindu/security-analyst
[35] www.forschung-it-sicherheit-kommunikationssysteme.de/projekte/mosaik
[36] www.forschung-it-sicherheit-kommunikationssysteme.de/projekte/sicia
[37] www.isograph.com/software/attacktree
[38] www.illinois.adsc.com.sg/cybersage/framework.html

4.1.3. Summary and Main Learnings

The current landscape shows a vast amount of different commercially available tools and approaches from the research and development domain for security in general and for security risk assessments in particular. In summary, two main learning can be described:

1. The categorised commercially available tools only cover certain ZCRs (1.1 / 2.1 / 5.2 / 5.13 / 6.1 - 6.9 / 7.1) from the IEC 62443-3-2 standard (as described within Section 3.6). In addition, the most resource-intensive ZCRs are not completely covered and not yet addressed adequately in an automated manner. The associated evaluation can be found within Section 5.1 and the corresponding scoping of the contents of this dissertation is shown in Section 5.2.

2. The achieved LOA ratings of the categorised commercial tools need to be further increased by additional research efforts. Currently, the maximum provided LOA value is 3 for the category of passive OT monitoring tools.

Furthermore, several additional aspects can be concluded from the available landscape of tools and approaches from the research and development domain. These are enhanced with insights from the available literature and are summarised within the following list [37, 44]:

- Only small focus on SMEs within the analysed tools and approaches
- Missing coupling with the safety domain
- Rarely well-documented implementations and real-world-based evaluations
- The creation of the information base is dependent on manual security expert knowledge elicitation
- Lack of conformity towards the globally-accepted IEC 62443 standard
- Information models do not follow standardised formatting rules
- Existing solutions are typically too extensive and complex
- Many statements and information about the commercial and proprietary tools cannot be checked in detail
- A solution should be independent from active or passive scans of networks and components
- Operator support through processed results and a GUI is necessary

The next section summarises the general security modelling basics which are needed to develop the expert system of this dissertation as a specific tool for the described problem statement (as stated in Section 1.3).

4.2. General Security Modelling

This section contains several parts each describing the state of the art regarding a certain type of formalism category for security modelling and the associated examples from the literature. In general, model-based formalisms can be used in a qualitative or quantitative manner to represent various aspects, e.g. system architecture, components, functionalities, or security properties [37, 152]. Whereas qualitative formalisms are highly favoured for security-related modelling in contrast to the safety domain which heavily relies on quantitative metrics [29]. Figure 4.1 shows an overview of available formalisms and which could be integrated into the developed expert system due to their relevance. All formalisms can be divided into three main categories: (1) **Mathematics** as formal, quantitative, and deterministic models, (2) **graphical / non-learning models** with semi-formal, either quantitative or qualitative, and partly fixed deterministic characteristics, and (3) **learning / reasoning** concepts as informal, either quantitative or qualitative, and typically indeterministic models. The formalisms of **Trees**, **Logic**, and **Class Models / Languages** highlighted in grey are directly impacting the specification of the expert system developed within this dissertation.

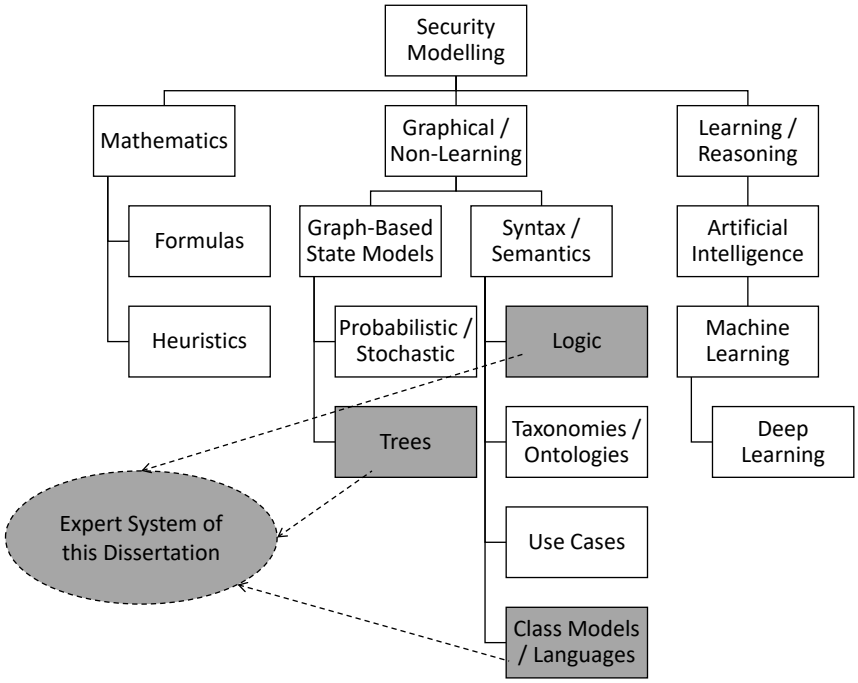

Figure 4.1: Overview of available security modelling formalisms

Each subsequent part includes a formalism from the second category of graphical and non-learning models due to their relevance for the initial semi-formal security modelling and the subsequent automation of security risk assessments in scope for this dissertation. Mathematics as formal models typically cannot be applied to larger and fuzzy problems without defined structures due to their complexity and dependencies [59]. The learning and reasoning formalisms as informal models are out

of scope due to a lack of transparency and credibility regarding possible results [61]. Each definition of a security modelling formalism follows a specific and repetitive structure including this information:

- Main characteristics of the group
- General advantages and disadvantages
- Associated approaches, concepts, and examples

The final discussion and evaluation of the formalisms is described later on in Chapter 7 regarding the usability, relevance, and alignment towards this dissertation.

4.2.1. Available Formalisms

Graph-based state models can be used for security-related attack and defense modelling and they consist of two main elements which make up the associated information model: (1) Nodes and (2) edges [153]. Nodes, also known as vertices, are used to model the concepts, e.g. events, goals, or actions. If the formalism is used in an inductive way, nodes can express causes (forward search) and if used in a deductive way, nodes can express consequences (backward search) [153]. The edges, also known as arcs, arrows, or lines, are the second main element of graphs. They connect nodes with each other and represent the relationship between them. In addition, edges can have special semantics to further enhance the description of the node relationships [153]. Generally, graphs can be described by their direction (directed vs. undirected) and their cyclicality (cyclic vs. acyclic) [153]. Nevertheless, graphs also have some disadvantages. The development of graphs is time-costly, especially when it is done for the first time for a given system [83], and is highly reliant on expert knowledge and available data [65].

In general, graphs are widely spread within the security domain and are extensively used for semi-formal information modelling due to their fitting characteristics and being able to model qualitative and quantitative contents [153]. Furthermore, there are also efforts to combine graphs with AI and ML approaches. Graphs are described as user-friendly and intuitive with the associated visual features. Therefore, graph-based security models are widely spread regarding a lot of different topics, e.g. IT or OT security and there specifically for the attack and defense modelling of SCADA systems or IACSs [153]. Additionally, the achieved result quality is typically adequate and the graphs are scalable to fit various sizes of information models [153].

With regard to security models there are many examples for the usage of graphs. The most prominent example are the Attack Graphs (AGs), also known as attack execution graphs [154], which typically model possible attacker behaviours and attack paths trying to exploit the SUC [55]. AGs are widely used and researched on. Therefore, there are already various tools available to automatically build them to represent a given system [154]. Other examples are privilege graphs [64], compromise graphs [64], or even combinations, such as the Bayesian Attack Graphs (BAGs) [55].

Another special type of directed graphs either cyclic or acyclic are the Petri nets typically used to model security characteristics of distributed systems [64]. They inherit great modelling capabilities for sequential aspects, concurrent actions, and various forms of dependencies [153]. In contrast to other graphical formalisms, e.g. UML or Business Process Model Notation (BPMN) diagrams, petri nets are provable by mathematics and therefore have additional formal modelling characteristics [64].

Trees are a special sub type of graphs, namely Directed Acyclic Graphs (DAGs) as a graphical and semi-formal state model, and are also based on nodes and edges to model hierarchical relationships [153]. As an enhancement of graphs, there are two types of nodes used within trees. Firstly, the root node represents the main event, goal, consequence, or action. It is the one designated node without any predecessor and it is connected with all other nodes via a dedicated path [153]. Secondly, the leaf nodes represent the primary or elementary events, goals, consequences, or actions. They are atomic components of a graph without any further refinement and do not have any children nodes [153]. Generally, trees enable a variety of usage, especially for all kinds of decision trees. They can represent very basic if-then statements or first-order logic (predicate logic) and can be enhanced with random forest approaches (multiple decision trees with aggregated result at the end) or gradient boosting (multiple decision trees with aggregated result on the way). In addition, trees can include quantitative and qualitative contents for analysis and are already widely used within the safety domain [83].

The most prominent examples are Attack Tree Analysis (ATA) to consider the paths an attacker might take, Attack Countermeasure Tree (ACT) to find out about possible defensive security improvements and investments, Fault Tree Analysis (FTA) already widely used within the safety domain, e.g. based on the IEC 13849 standard, and Event Tree Analysis (ETA) to identify unexpected system conditions within the industrial domain [64, 65, 83]. Other examples are trees to model vulnerabilities, attack responses, compromises, insecurity flows, intrusion, security activities, or the Bow-Tie analysis [65, 153, 155]. In addition, trees can be used to model game theory or red teaming approaches representing the interaction between attackers and defenders as a non-cooperative, sequential, perfect information, and non-zero sum game [64, 65]. Also the formalisms following the BPMN, such as the lightweight swimlanes, fall under the tree category.

Probabilistic and stochastic formalisms are graphical and non-learning models for probability distributions and relations over a set of random variables. For the security engineering domain, these formalisms are typically used under the scope of Probabilistic Risk Assessment (PRA) methods which are the de facto standard for safety risk assessments regarding critical facilities [65]. They are very popular among researchers and practitioners due to the convenient and easily understandable quantitative estimation of risks. The gained insights can then be used to assist decision makers to better understand the current and the target security levels of an organisation [64]. The general strength of these formalisms lies within the scenario-based approach that systematises knowledge and uncertainties about the SUC to include malicious activity into the overall security risk assessment processes [65].

The PRA is then typically used for Probabilistic Safety Assessments (PSAs) and Quantitative Risk Assessments (QRAs) [65] coming from safety and security domain alike.

PRA methods can be used to formalise and analyse security under uncertainty [153]. These formalisms are flexible risk assessment tools and offer the potential to reduce the needed expert knowledge and the amount of unidentified risk scenarios [65]. Nevertheless, there is no way to deal with unknown vulnerabilities, techniques, or attackers and their effect on the risk estimation making it incomplete in the pure mathematical sense [64]. PRA methods either need historical data or rely on subjective data, which typically have a limited availability and therefore inherit a reduced applicability in general [64]. Furthermore, these formalisms are based on simplifying assumptions and do not cover risks with low probabilities in an adequate way [64]. The group of probabilistic and stochastic formalisms includes various approaches and concepts. A short summarising overview is provided here.

Bayesian Networks (BNs), also known as belief or causal networks, are directed graphs used to model probabilistic influences for security based on quantitative results [55, 65, 83, 153, 156]. They are always dependent on training data to build up the network of nodes representing events associated with probabilistic variables and directed edges representing causal dependencies between the nodes [153]. The general aim is to use probabilistic inference techniques to perform security risk assessments based on the predicted likelihood of future attacks [153]. Another term for this kind of formalism are the Bayesian Belief Networks (BBNs) as probabilistic graphical models which are already used for safety and security risk assessments [83, 157].

Linked to the BNs are the Monte Carlo simulations which are used for a consideration of all elements within a system in a random manner to explore unknown risk scenarios [55]. Monte Carlo simulations can be used to feed BNs, but are typically resource-intensive and require long run times [65].

Markov models, also known as Markov networks, chains, or random fields, are undirected graphs used to model probabilistic influences without a clear directionality and can also be integrated into BN analyses [55, 65]. In general, there are four types depending on the characteristics of state space (discrete vs. continuous) and the time variable (discrete vs. continuous). They can be used to model upcoming events independent of the past [65]. Furthermore, similar to Monte Carlo simulations it is possible to reduce the set of unknown risks by using Markov models. Special applications for the security domain can be also based on Continuous Time Markov Chains (CTMCs), Discrete Time Markov Chains (DTMCs), Variable Length Markov Chains (VLMCs), or Semi-Markov-Chains [154]. Using Markov models and combining them with fault trees [64], Boolean logic Driven Markov Processes (BDMP) models can generate qualitative and quantitative outcomes for safety and security risk assessments respecting the real-time characteristics of industrial applications [84].

Syntax & Semantic are the basis for our everyday communication in, e.g. human conversations, written languages, or software coding. They are typically implemented by graphical or linguistic elements in a human-readable format [59]. Thereby, syntax comprises all the rules and regulations to be grammatically correct ("how") and

semantic describes the associated meaning and content ("what"). The usage of these basic principles for the information modelling of security enables the creation of semi-formal models, which are not mathematically provable, but can be used to build the basis for standardised understandability and interoperability [59, 158]. The corresponding information modelling is a major research challenge regarding engineering of industrial systems due to the wide range of application possibilities.

Typical formalisms from this category are logic-based approaches, taxonomies, ontologies, use case diagrams of various kinds, and class models including modelling languages [159]. They try to capture expert knowledge in a graphical and structured way providing the following benefits: Security experts can distribute their knowledge in an easier way, security novices can use the formalised expert knowledge to tackle issues on their own, and the current security challenges can be solved in a structured way [59]. The only disadvantage is the necessity to create these semi-formal information models manually depending on the respective expert knowledge, which makes comparison, interoperability, and automated validation more difficult [59].

The domain of logic offers a variety of approaches to formalise expert knowledge. Examples are the propositional logic, predicate logic, temporal logic, or fuzzy logic [58]. These formalisms can be used to model rule-based expert systems and can be enhanced with semantic queries based on SPARQL Protocol and RDF Query Language (SPARQL) regarding inference or rule checks of the syntax based on Shapes Constraint Language (SHACL) or Semantic Web Rule Language (SWRL) which are called reasoners [44]. Nevertheless, the queries and rule checks require standardised and agreed modelling of information and the performance is highly dependant on the structure of the model.

The next type of semi-formal and graphical models are structured objects within taxonomies and ontologies [160]. A taxonomy is used to categorise and organise a certain domain of information based on discrete sets following a hierarchical schema with an abstract point of view [161, 162]. Humans use taxonomies on a daily basis without specifically recognising them as those, e.g. when identifying animals or plants. Ontologies are specified in a more detailed way including relationships between the modelled objects to show the properties of a subject area [163]. Therefore, ontologies are designed in a more low-level approach in comparison to the high-level taxonomies. There are various approaches and concepts available to model taxonomies and ontologies for different applications, e.g. Web Ontology Language (OWL), Turtle, or the Resource Description Framework (RDF) with the Resource Description Framework Schema (RDFS) [44, 55, 146, 164, 165]. These formalisms use either an open world assumption (information not modelled = no statement possible) or a closed world assumption (information not modelled = invalid) [44]. In addition, reasoning or inference mechanisms can be used to deduct new knowledge from these models or to check for consistency of models [58]. Ontologies are based on logic rules ("axioms") and combine the approaches of the semantic networks and frames resulting in a higher expressive power [44]. When taxonomies or ontologies are used to define a certain domain with specific knowledge for a dedicated use case or application the result is called a Knowledge Graph (KG) [166]. By doing so, the generally valid statements

are translated from the terminological level containing types and classes into the assertional level representing specific individual manifestations or instances [44]. Various examples from the security domain are available proving the usage of graphical approaches, e.g. the Cyber Terrorism SCADA Risk Framework [64], vulnerability analysis for Cyber-Physical Systems (CPSs) [167], automated requirements management [168], attack path analysis [169], the Semantic Processing of Security Event Streams (SEPSES) KG [166], the C2NET platform for decision-making [170], risk assessment for the IT domain based on the Common Attack Pattern Enumeration and Classification (CAPEC) methodology [45], the Security Domain Specific Visual Language (SecDSVL) to support enterprise security modelling [171], or a combination of safety and security into one common ontology [172].

Class models and the associated modelling languages comprise an additional type of graphical information models. They are typically used to formalise characteristics of certain objects and their interrelations between each other [59]. The most prominent example is the UML with various types of diagrams available and the corresponding extensions especially for the security domain [173, 174]:

- SecureUML [114, 146]
- UMLsec [114, 146]
- SysML and SysML-Sec [55, 84, 114, 146, 175]
- IoTsec [146]
- Automation Markup Language (AML) [55, 146]
- AMLsec [32, 146]
- Model-Based System Engineering (MBSE) [176, 177]
- Security Requirements Engineering (SRE) [149]

Similar to class models, various diagram types are available to model human behaviour regarding security-related contents, e.g. attack or defense. This category of formalisms includes use case, misuse case, activity, or abuse diagrams for sequencing events and the associated Data Flow Diagrams (DFDs) [55, 113, 114, 168]. These formalisms are used to describe the relationships within a complete system [59]. Their flexibility and intuitiveness allows expressive semi-formal and graphical information modelling of, e.g. attack scenarios, defensive mechanisms, or human behaviour, without the need for a mathematical formalisation [153].

4.2.2. Transition to Expert Systems

Expert systems are computer programs that emulate human experts from a specific domain regarding their knowledge-related capabilities to solve complex problems [44,

178]. They can be categorised as knowledge-based systems like illustrated by Figure 4.2 in alignment to the used definition of this dissertation. Accordingly, expert systems are special knowledge-based systems where the information are elicitated from a human expert who has qualitatively high knowledge regarding content, quantity, abstraction capabilities, and reasoning [178]. Typical experts have above-average abilities to solve their domain-specific problems, often act intuitively right, can work with incomplete data or uncertainties, are rare and expensive, have long-term experiences and learnings, and their knowledge cannot be passed on easily [178]. The main aspect of an expert system is the separation between the representation of knowledge within a knowledge base and the usage of this knowledge within a knowledge processing [178]. This allows a clear differentiation between information modelling for the problem description and the problem solving itself [44].

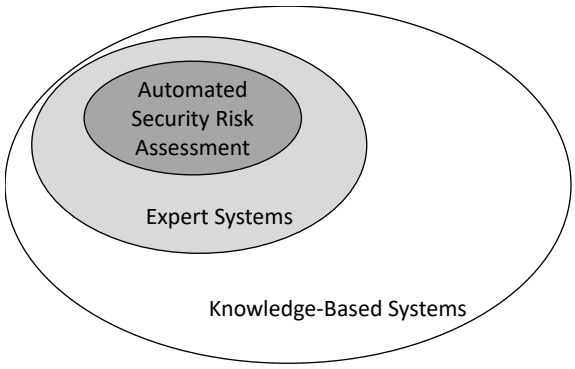

Figure 4.2: Categorisation of knowledge-based and expert systems for security risk assessments in alignment to this dissertation (adapted from [44])

In general, there is a huge potential to automate certain steps regarding security risk assessments by using expert systems [44]. Especially, SMEs can benefit from the provided OT security knowledge and will be able to reduce their overall costs and personnel efforts to perform security risk assessments [44]. During the course of digitalisation, new industrial systems are emerging combining social and technological aspects which can support the human experts for typical assessment, evaluation, and audit tasks [33]. Typically, results from computer-based expert systems are cheaper when used for repetitive tasks [33]. The work of [44] estimates support possibilities for knowledge-based systems regarding the automation of security-related tasks on a qualitative scale ranging from very low to very high. Table 4.1 summarises these estimations in alignment with the defined ZCRs from the IEC 62443 standard for this dissertation.

In general, expert systems can be used in various domains, e.g. software development, architecture, logistics, power, robotics, medicine, engineering, language processing, and management [60]. Furthermore, they are able to understand and solve problems, to explain and evaluate solutions, and to acquire and to structure the underlying knowledge [60]. This results in a high complexity of the expert system

Table 4.1: Support possibilities for knowledge-based systems regarding the automation of general security-related analysis process steps (adapted from [44])

Process Step	**Support Possibility**	**ZCR Alignment**
Preparation	Low	1.1
Target Asset Identification	High	1.1 & 2.1
Threat Identification	Very High	5.1
Countermeasure Selection	Medium	5.8 & 5.12
Countermeasure Implementation	Low	-
Documentation	High	5.13 / 6.1 - 6.9 / 7.1

itself which needs to be reduced and abstracted towards the operator, but also offers the possibility to achieve a high reliability, transparency, and credibility of the results, especially needed decision making [60]. In addition, expert systems are easily maintainable and expandable, deliver justified and explainable results, enhance productivity and availability, and reduce downtimes and decision-making times [178]. In contrast to the overall predominant advantages, expert systems are hard to create and obtain, require the dedicated input from human experts, and are mostly only suitable for a certain range of tasks under specific conditions [60, 178].

The area of expert systems also includes additional functionality and characteristics. In general, it can be distinguished between an open world assumption (information not modelled = no statement possible) and a closed world assumption (information not modelled = invalid) in alignment towards the definitions surrounding taxonomies and ontologies mentioned before [44]. The security domain is typically modelled as an open world assumption due to, e.g. not yet discovered or published technical vulnerabilities or unknown threat actors, tactics and techniques [44]. Furthermore, advanced expert systems offer the abilities of resolution and inference. Inferencing new knowledge is based on the creation of new conclusions out of a present set of premises [44]. There are two types of inference available: (1) Data-driven inference (called forward-chained) with knowledge as the starting point and (2) goal-oriented inference (called backward-chained) with conclusions as the starting point [178]. In addition, resolution describes the process of evaluating rules within an expert system regarding their validity by indicating a logical contradiction [44].

4.2.3. Summary and Expert System Definition

The presented status of available security modelling formalisms (also see Figure 4.1) paves the way towards the desired goal of this dissertation represented by a specific expert system utilised for the automation of security risk assessments within the OT domain [142]. In theory, all of the mentioned security modelling formalisms (from Section 4.2.1) are adequate to be used as the foundation of an individual expert system. But following the previous work by [44], it appears that forward-chained (starting with asset, vulnerability, and threat information in a tree structure as

described in Section 6 and Section 7) and rule-based (using a set of rules as specified in Section 8) expert systems are the best choice to support SMEs with automated security risk assessments.

The necessary knowledge base for the expert system is built upon the usage of DTs in the form of the AAS following standardised class models and language formats which is further shown in Section 9 and the knowledge processing is done via a prototypical implementation illustrated in Section 10. To ensure a high credibility, comprehensibility, and traceability with regard to the additional safety requirements, no inference functionalities are used [61]. Consequently, the ability to check newly created rules via resolution is not necessary due to only manually created logical rules and the absence of learning-based concepts which might result in new and untested knowledge within the expert system. Nevertheless, the descriptive term of forward-chained can be used to describe style of the desired expert system of this dissertation. In addition, it works in an adapted manner of a closed world assumption with certain constraints based on a finite set of information which is not automatically expandable [65] and guarantees the absence of uncertainties [58]. This leads to the following security modelling definition for the expert system regarding automated security risk assessments of this dissertation:

Forward-chained rule-based expert system using a fixed knowledge base (deterministic information model) and a scaled-down knowledge processing (predicate logic rules) without reasoning and inference capabilities for a practical implementation of an automated security risk assessment achieving high credibility and traceability.

By further narrowing down the scope of contents and the focus of this dissertation based on the given definition, it is possible to identify, analyse, compare, and evaluate the related work as part of the overall state of the art. The summary of results is presented within the next section.

4.3. Related Work: Automation of Security Risk Assessments

After identifying and presenting the available commercial tools and approaches for security risk assessments and the general security modelling formalisms, this section summarises, compares, and evaluates the related work of techniques for this dissertation based on a defined set of security risk assessment characteristics. The section builds upon previous works by the author of this dissertation and extends the published results from [56, 57, 179]. The final results are later shown in Table 4.2 within Section 4.3.2. The first two columns contain the identifier and the referenced publications for further interest regarding the source of information. The last column describes the overall relevance of the specific related work on a three-tier scale (low - medium - high) towards this dissertation based on a subjective evaluation. If certain

characteristics are not fitting to the specific related work, the table cell is marked with "n/a" representing not available information. The rest of the characteristics used within Table 4.2 is described in the following Section 4.3.1 in a more detailed way including the possible options.

4.3.1. Identification of Characteristics

The **domain** includes the originally intended scope of the proposed security risk assessment technique sorted by the following choices [2]:

1. OT, e.g. IACS, ICS, or SCADA
2. IT, e.g. Information and Communication Technology (ICT) or office-related
3. Other, e.g. critical infrastructures, energy, process industry, or transportation

Completeness describes the process coverage of the security risk assessment technique referring to the extent in which the results resemble reality. The five typical phases are [150, 180]:

1. Preparation and scoping
2. Risk identification (to find, identify, and describe risks that could help or hinder an organisation to achieve its objectives)
3. Risk analysis (to understand the nature of the risk, its characteristics, and where appropriate, the level of risk)
4. Risk evaluation (to support decisions by comparing the results of the risk analysis with the established risk criteria to determine countermeasures)
5. Documentation

The **goal** characteristic contains the intended aim of the security risk assessment technique following an abstract scale independent of the associated domain. The respective results can then be integrated into the typical risk management decisions [2]:

1. Identify need for controls
2. Cyber-informed safety analysis
3. Prioritise controls
4. Identify optimal strategy

The characteristic of the **starting basis** specifies the origin of the security risk assessment technique and from which viewpoint it is performed. There are four main types distinguished and hybrid combinations are also possible. Options are [2]:

1. Threat
2. Asset
3. Impact
4. Vulnerability
5. Impact and threat
6. Asset and threat
7. Asset and vulnerability
8. Threat and vulnerability
9. Asset, threat, and vulnerability

The **data source** column includes the possible inputs for the security risk assessment technique. The following list shows some examples of acknowledged and accepted technologies, approaches, frameworks, or other sources of information in contrast to a manual security expert knowledge elicitation [2]:

1. Vulnerability data, e.g. CVE, Common Vulnerability Scoring System (CVSS), Common Weakness Enumeration (CWE), or Common Weakness Scoring System (CWSS)
2. Threat and attacker modelling, e.g. MITRE ATT&CK framework, CAPEC, or Intel Threat Agency Library (TAL)
3. Countermeasures, e.g. MITRE ATT&CK framework or MITRE D3FEND
4. Manual security expert knowledge elicitation

The constantly changing environment and requirements faced by security risk assessment techniques make the **repeatability** a main factor for traceability. It is also important to produce the same results from different operators [2, 180]:

1. No capability for repeatability
2. Low capability for repeatability (reliance on qualitative data and/or experts)
3. Medium capability for repeatability (combination of sources)
4. High capability for repeatability (trusted data sources)

Security risk assessment results need to be available in a continuous, adaptive, and efficient way. The **timing** characteristic summarises these characteristics [154, 170]:

1. Static, e.g. only initially, once, or on fixed points in time
2. Dynamic, e.g. real-time, during run-time, or triggered

The **perspective** includes the resource valuation and the view on the respective SUC. The four choices are [48, 142, 175]:

1. Asset-driven
2. Service-driven
3. Quality-driven
4. Business-driven

Safety coupling is essential for a future-proof security risk assessment technique (as already discussed in Section 3.4). The literature proposes four types of relationships [31, 77, 83]:

1. Complete antagonism: When considered jointly, safety and security requirements or measures lead to conflicting situations
2. Independence: No interaction at all
3. Conditional dependency: Fulfilment of safety requirements conditions security or vice-versa
4. Mutual reinforcement: Fulfilment of safety requirements or safety measures contributes to security, or vice-versa, thereby enabling resource optimisation and cost reduction

The **assessment type** describes how the results of the security risk assessment technique can be characterised. This is an essential factor for the appraisement of results and their integration into the overall risk management processes. There are three types [2, 48, 64, 150, 154]:

1. Quantitative
2. Semi-quantitative (hybrid)
3. Qualitative

The standard **compliance** of a security risk assessment technique is important to achieve the needed credibility and trust. In addition, it is required to reach a certain conformity with the state of the art and to be assured for a future usage [150, 154, 181]. The column is filled with free text including a list of the standards and best practices addressed by the respective related work.

Each security risk assessment technique has an own **adoptability**. This characteristic describes the implementation status based on the TRLs and the level of rigor representing the difficulty, time, and resources required to adopt the respective proposed solution approach [2, 64, 150, 154, 181]:

- Research → TRLs 1, 2 & 3 (high level of rigor)
- Development → TRLs 4, 5 & 6 (medium level of rigor)
- Deployment → TRLs 7, 8 & 9 (low level of rigor)

4.3.2. Comparison of Techniques

Table 4.2: Overview, comparison, and evaluation of the related work based on the previously defined security risk assessment characteristics

Identifier	Reference	Domain	Completeness	Goal	Starting Basis	Data Source	Repeatability	Timing	Perspective	Safety Coupling	Assessment Type	Compliance	Adoptability	Relevance
#1	[146, 175]	OT	1, 2 & 5	1	9	1 & 5	3	1	3	2	n/a	IEC 62443-3-2	Development	High
#2	[44]	OT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	n/a	IEC 62443 ISO/IEC 2700x VDI/VDE 2182	Research	High
#3	NAMUR WG 1.3 [32, 47, 128, 182–184]	OT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	2	n/a	Method-agnostic	Research	High
#4	[185]	OT	1-4	3	9	1, 2 & 3	3	2	1	3	3	IEC 61508	Development	High
#5	[46]	OT	1-5	1	6	4	4	2	1	2	n/a	IEC 62443	Research	High
#6	[87]	OT	2-4	2	2	4	2	1	1	3	n/a	IEC 62061	Development	Medium
#7	[51]	IT	1-3	1	7	1 & 4	3	2	1	2	1	ISO/ISO 2700x	Research	Medium
#8	[186]	IT	1-4	1 & 3	9	4	2	1	1	2	1	NIST SP 800-30	Research	Medium
#9	[187]	Other	1 & 5	4	2	4	3	1	1	2	1	NIST SP 800-53	Research	Medium
#10	[180]	OT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	2	n/a	IEC 62443 ISO/IEC 2700x NIST 800-53	Research	Medium
#11	[120]	OT	n/a	n/a	n/a	n/a	n/a	n/a	n/a	2	n/a	IEC 62443 VDI/VDE 2182 ICS CERT	Research	Medium
#12	[188]	OT	3	3	1	2	4	1	n/a	2	1	n/a	Research	Medium
#13	[189]	OT	1 & 2	1 & 3	9	1 & 2	3	1	1	2	n/a	n/a	Research	Low
#14	[145, 151]	OT	1-3	1	6	4	2	1	2	2	n/a	n/a	Research	Low
#15	[74]	OT	1	n/a	8	1 & 2	4	2	1	2	n/a	IEC 62443	Research	Low
#16	[78]	OT	1	1	n/a	4	2	2	n/a	4	n/a	IEC 62443	Development	Low

#1: In [146] a tool for risk identification is developed for OT systems during the engineering phase of the lifecycle. This is based on the modelling of the system structure (ZCR 1.1) as well as asset information in AML which is then translated into an OWL ontology using a preliminary approach from [190]. The self-developed enhancement "AMLsec" is used as a knowledge base and enables semantic inference and reasoning to create and test new knowledge. Afterwards, queries and rules written in SPARQL and SHACL are used to identify threats, vulnerabilities, and consequences (ZCRs 5.1 - 5.3) automatically to comply with security requirements (ZCRs 3.2 - 3.6) from the IEC 62443-3-2 standard. By doing so, attack graphs are generated and pruned afterwards to show the most critical paths for possible attackers. The implementation[39] is evaluated using a pre-defined and exemplary AML case study of a robot cell[40] regarding performance and scalability. A security-

[39] www.github.com/sbaresearch/amlsec
[40] www.automationml.org/news/example-file-of-a-robotcell

or content-related evaluation is missing. The presented tool offers well-developed and security-related capabilities for risk identification in the form of a compliance check, but lacks the risk analysis and risk evaluation. Also the integration of DTs is missing and the maintenance of the required ontology seems to be too complex. In addition, safety is only used as a criteria for network segmentation and not to affect the security risk assessment process itself.

#1: The work of [175] further extends and develops the results from [146]. Within the update, the QualSec method is proposed which can be used for the automated security risk identification for Cyber-Physical Production Systems (CPPSs) during the engineering phase. Therefore, the quality characteristics of products are assessed as a basis for the risks in contrast to the typical asset-centric view. This method is based on a semantic representation of engineering knowledge specified within AML models, especially plant topologies from Computer Aided Engineering Exchange (CAEX) and sequencing information from PLCopen Extensible Markup Language (XML). In addition, the QualSec method creates quality-driven petri nets to perform the security risk identification. Both publications ([146] and [175]) contain very sophisticated approaches for the automation of security risk identification within the OT domain using knowledge- and rule-based concepts. In contrast, this dissertation aims at providing an automation of the complete risk assessment process including risk identification, risk analysis, and risk evaluation. Nevertheless, the research results in [146] and [175] represents the related work with the highest relevance for this dissertation.

#2: In [44] a general knowledge management concept is developed that ensures the timeliness and correctness of OT security knowledge and provides it in the required scope and level of detail depending on the lifecycle phase of a production plant. The concept is based on four main parts consisting of a knowledge model, an OT security knowledge lifecycle, knowledge management, and the overall integration. An implementation and an associated practical testing are missing. Therefore, the work remains on a conceptual and general level. Nevertheless, [44] represents an important part of the overall related work covering side topics of this dissertation and paving the way for a consistent knowledge management within the OT domain also affecting the goal of automating the security risk assessment process. Especially, the topics of knowledge-based systems, the rule-based definition of expert knowledge, the introduction and motivation, and the provided state of the art are highly relevant for this dissertation.

#3: Another highly relevant related work is still under development within the security engineering subgroup associated with the NAMUR 1.3 working group[41]. There, a generally-applicable and method-agnostic reference model for security engineering independent of the lifecycle is specified to establish a common language and understanding towards the various engineering disciplines outside of security.

[41]The author of this dissertation is already participating within the working group to synchronise on the related topics and to exchange research and knowledge insights. Link: www.namur.net/en/work-areas-and-project-groups/wa-1-project-planning-and-construction/wg-13-information-management-and-tools.html

The reference model was already presented at the NAMUR general assembly in 2022, is in an internal review phase to be transformed into the NAMUR recommendation NE 193 during the time of writing this dissertation, and is prepared to be integrated into a security engineering AAS submodel[42]. The NAMUR reference model is influenced by the works from [47, 128, 182, 184] proposing a method-agnostic and generally-applicable thought model for security engineering and the associated visualisation of security design decisions. This happens in close alignment with the German research project "Integrated Data Models for the Engineering of Automation Security" (IDEAS)[43]. Figure 4.3 summarises the research focus of this dissertation (in grey) and the alignment towards the NAMUR reference model. The part regarding security risks defined within the NAMUR reference model is detailed by a use-case-oriented specialisation and an implementation for the automation of security risk assessments regarding modular IACSs. A further comparison of contents is not yet possible due to ongoing activities, but the basic ideas are already discussed and integrated into the NAMUR subgroup to benefit from the research and information exchange. In [32] and [183], results from the same concept are published showing first steps to combine general engineering and security workflows in a machine-readable way as a digital enabler for security by design. This includes the necessary definition of use cases and requirements for a security engineering information model and the integration into AML. See Section 9.2 for further details on the context of AML and the AAS.

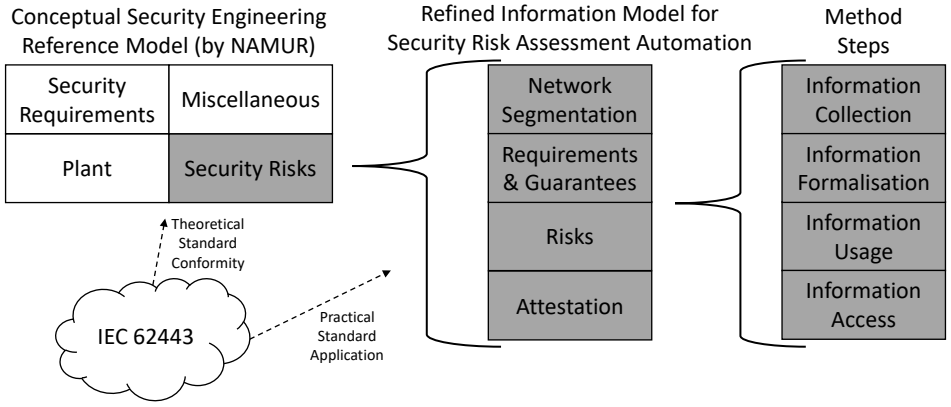

Figure 4.3: Research focus and alignment of this dissertation towards the presented reference model from the NAMUR 1.3 working group

#4: The authors of [185] propose a network-based IDS for OT environments being capable of coupling safety and security. The detected security-based incidents are used to evaluate possible safety-related impacts on the SUC. Afterwards, the information are used for a risk management strategy to prioritise attacks and to suggest adequate countermeasures based on a self-defined severity of MITRE ATT&CK tactics and

[42]www.industrialdigitaltwin.org/en/content-hub/submodels
[43]www.forschung-it-sicherheit-kommunikationssysteme.de/projekte/ideas

the probabilities gained from the safety-related PL and SIL values of the affected assets. The proof of concept utilises an anomaly-based approach via a random forest for the learning-based traffic classification as a basis for the IDS and a small subset of typical OT attack vectors. The work of [185] represents a very good example of security-related technologies coupled with the safety domain to achieve a common result. Nevertheless, the implementation is not completely comprehensible and the approach of a network scanner to gain data for the IDS is seen as not adequate for a non-intrusive OT security risk assessment as intended for this dissertation.

#5: In [46], a partly automated security analysis approach based on a rule- and knowledge-based system is presented. The shown tool is able to import system information via CAEX files, to perform a threat analysis on the scoped SUC, and to evaluate possible countermeasures addressing identified risks. Engineering artefacts specified in AML enhanced with OWL and SWRL are used to discover security flaws and vulnerabilities during the engineering phase of manufacturing systems. By doing so, a time and resource reduction is possible. Nevertheless, the results are just presented on an abstract way and it is not exactly comprehensible in which way and how much reduction is achieved. In addition, tedious manual tasks, such as the definition of a target (represented by an SL-T from the IEC 62443) and the maintenance of the expert knowledge, remain and a coupling with the safety domain is only proposed as future work.

#6: The authors of [87] propose the Safety & Security Combination (SafeSecCombi) approach to couple safety and security concerns for automating digital production plants in alignment with the IEC 62061 requirements. It can validate effects on safety, identify assets causing unsafe behaviour, and analyse security risks, which results in improved analysis capabilities in contrast to separated safety and security risk assessments. The SafeSecCombi requires a previously performed and correctly documented Failure Modes and Effects Analysis (FMEA) and a model of the production environment based on Product-Process-Resources (PPRs). The approach needs to be used by a team comprised of different engineering roles, e.g. domain experts, quality managers, safety experts, or security experts, and offers the possibility to cover all risk assessment phases. Nevertheless, the current state of research shows a high need for manual or partly semi-automated preparation of information, e.g. the FMEA, the PPR model, or the Production Asset Network (PAN). In addition, based on the current evaluation use case the approach is intended for the automotive production [87].

#7: Within the work of [51] from 2007, a model-based security risk analysis is presented and it is shown how a certain degree of automation can be achieved. XML is used to describe the IT assets, systems, and network architecture. The security risk analysis itself consists of two aspects: (1) Evaluation of available vulnerability information of the assets and (2) compliance checks regarding defined best practices, e.g. present firewall at network entry points. At the end, a two-dimensional lookup table is specified to assess the final security risk value.

#8: A similar work for the IT domain was proposed by [186] in 2009. There, the risk management framework AURUM based on the NIST SP 800-30 is described and

evaluated by a comparison with the CRISAM tool[44] and the former BSI GSTOOL. The ontology-based methodology is able to determine likelihoods and impacts for security risks based on system, threat, and vulnerability information to recommend and select the appropriate countermeasures. The presented approaches utilise similar concepts and data sources, such as threats, technical vulnerabilities described via CVE and CVSS, as compared to this dissertation. Unfortunately, the implementations are not thoroughly described and therefore not comprehensible. In addition, due to intended use for the IT domain a coupling towards safety and the adaptation for the OT domain are completely missing.

#9: In [187], a tool written in Prolog for the automated assessment of security compliance checks of critical infrastructure, especially electrical power grids, is presented. The system architecture can be integrated into a dependency graph via a cyber infrastructure modelling language and predicate logic in the form of horn clauses is used to define the rules deducted from standards and best practices. In addition, a visualisation of the system architecture and the compliance check results is given to support decision-makers. Furthermore, safety is not regarded at all.

#10: In [180], requirements, typical contents, types, and sources of knowledge for ICS security analyses are derived. The result is an ontology on a conceptual level consisting of four layers and a knowledge lifecycle fulfilling the defined needs and joining existing works from various domains into one proposal. The integration into tools for knowledge management and ICS security analyses are missing, but the specified ontology can be used as a knowledge-based system, e.g. for a rule-based inference of information. The defined aspects for security knowledge are aligned with the contents of this dissertation and are used as a basis for the definition of the information model. Nevertheless, the topic of automating ICS security analyses is not covered in [180], neither the associated security risk assessment processes.

#11: In [120], a security risk identification approach (even though the wording within the publication uses the term security-related diagnosis) is presented based on a self-developed categorisation of industrial assets. Engineering artefacts specified within AML enhanced with an OWL model are used to detect security flaws during the engineering phase of manufacturing systems. The publication contains a structured information base for the automated security risk assessment, but also pointing at the fragmentation of information sources and demanding a unified, vendor-neutral data format. This dissertation adapts contents from [120], such as sources and categories of system information, and integrated them into the presented solution approach with a clear focus on practical security risk assessment processes and on the creation of a DT template in the form of the AAS to achieve a common and standardised information model in contrast to the proprietary AML modelling.

#12: In [188], the quantitative and probabilistic ATT&CK-OAHP-BN security risk assessment method for information systems is proposed. The authors create a knowledge graph based on the MITRE ATT&CK framework with the help of the neo4j graph database tool. An Optimized Analytic Hierarchy Process (OAHP) identifies

[44] www.crisam.net

the weight of the security indicators according to a hierarchical structure model and a BN then calculates the probabilities of attacker techniques to be protected by the system designer. The Dragonfly adversary group is chosen here as an example. The presented approach only covers the security risk analysis phase and does not regard safety at all. In addition, the quantitative probability-based definition of security risks lacks a real-world correspondence and proofs.

#13: A similar approach is used in [189] originating from the EU research project INSPIRE[45] which started in 2008. A self-developed security ontology containing assets, safeguards, threats, attack sources, and vulnerabilities is used for reasoning via SWRL rules. The result is a decision aid tool including vulnerability and threat identification for SCADA components within the critical infrastructure domain. Unfortunately, the solution approach is only described on a very abstract level and details are missing. Nevertheless, this is one of the first works describing automated security analyses for the OT domain.

#14: In [145] and [151], an approach for repetitive security analyses regarding I4.0 systems is presented. The proposed tool is implemented in Java using the Eclipse framework and utilises MBSE principles for a Threat And Risk Analysis (TARA) based on attack trees. Pre-defined libraries of security requirements with associated countermeasures and threats are provided within the tool to support the still manual process of specifying the system architecture via UML models, choosing threats, and determining their impact and feasibility. Based on the manual inputs security countermeasures are proposed for further deployment. Unfortunately, the specification and creation of the used libraries is not further explained and therefore not comprehensible. In addition, the automation degree is quite low in comparison with other works resulting in a high rigor for a possible repetition.

#15: The work by [74] proposes a twofold ontology-based approach for automated and minimally-invasive (1) security analyses and (2) incident responses. For this dissertation only the first part regarding the security analyses is important and further regraded here. This includes the topics of configuration, conformity, vulnerability, and threat analysis plus attack detection and correlation. The developed, prototyped, and evaluated System Model Processing (SyMP) approach is a consistent separation-of-concerns-based framework for knowledge-based security analyses. The focus is set on the semantic interplay of various ontologies and the deduction of knowledge via predefined rules based on statements from typical OT standards and guidelines. The approach is very sophisticated in terms of ontology handling and usage, but unfortunately only covers certain parts of of security risk assessments, e.g. vulnerability or threat analysis.

#16: In [78], a coupled approach for safety and security risk assessments based on a commonly shared ontology is presented. The ontology is created with OWL using the online version of the Protégé tool[46]. In addition, SPARQL queries are specified and used to infer knowledge from the ontology regarding relations, such as

[45] www.cordis.europa.eu/project/id/225553
[46] www.protege.stanford.edu

hazards, incidents, and threat events. The overall proposal is tested and evaluated using a small research-grade modular production system demonstrator from Festo[47]. Nevertheless, the publication does not contain any contents regarding the real assessment of safety and security risks. The provided concept can be used to base a security risk assessment, but additional research work needs to be included here to reach a certain level of usability and automation.

4.3.3. Summary and Definition of Deficits

The overall status of the related work for automated security risk assessments shows a strong coverage of security-related approaches, tools, and techniques for the lifecycle phases of development and engineering of OT systems, but also for various other domains, such as IT, critical infrastructure, energy, or transportation (see Section 4.1). The full range of available security modelling formalisms from Section 4.2 is utilised within the related work to achieve improvements regarding security risk assessments, e.g. by rule-based systems, taxonomies or ontologies, ML and AI, or mathematical formulas. In addition, there are various works available which only introduce an automation of certain steps, e.g. attack path generation for security risk identification. Nevertheless, several aspects for an automation of security risk assessments are missing, e.g. the automated risk analysis and risk evaluation, the support during runtime and operation of an OT system for asset owners and system integrators, the flexibility required for modular manufacturing systems, the coupling towards the safety domain, and the conformity to the IEC 62443. Furthermore, there is a general lack of research regarding assessing and documenting security for the OT domain [49].

The complete analysis of the state of the art including all the mentioned aspects leads to the identification and definition of four main deficits which are also already stated during the problem statement in Section 1.3 as a preview. They are numbered via an identifier and explained in the following:

- **D1: Insufficient process coverage** - The complete security risk assessment process is not covered, e.g. risk analysis and risk evaluation are missing which results in a lack of final risk determination. In addition, the coupling with safety processes, models, or objectives is not regarded within the related work.
- **D2: No standardised metrics** - Typically, proprietary formalisations are proposed that contrast well-established security metrics and block the reusage of knowledge. Moreover, the general conformity towards globally-accepted standards is missing. A lack of details and poor user guidances hinders the understanding of the approaches. In addition, the usage of subjective data limits the comprehensibility.
- **D3: Low maturity** - There are rarely implemented and evaluated approaches available. The maximum TRL is typically 4. Furthermore, the improvement

[47] ip.festo-didactic.com/Infoportal/MPS/MPS403I4.0/EN/

of the automation degree is often not clear in terms of resource savings or knowledge formalisation and the security improvement is therefore not measurable. In many cases a sophisticated GUI as interface and improvement of the usability for the operator is missing as well.

- **D4: High abstraction level** - The proposed models are typically not suitable for real-world scenarios due to missing interfaces towards assets or a generic concept level. Furthermore, most of the models are not directly applicable and need additional manual refinement for the targeted use cases.

The remaining sections are build around the identified state of the art and the evaluated related work building the foundation and the guiding framework for this dissertation. Each of the four deficits will be answered within one of the following method steps for information and process modelling as described within the following Sections 6, 7, 8, and 9.

Part III.

Method Steps for Information and Process Modelling

5. Overview and Scoping

This chapter is based on the authored and already published publication of [53] and comprises the overview of the procedure focusing on the scoping of the conducted four main method steps. This includes the practical analysis of the processes through three separate manual security risk assessments performed in practical case studies shown in Section 5.1. Afterwards, the identified focus for the automation of security risk assessments is presented in Section 5.2.

5.1. Security Risk Assessment Case Studies

5.1.1. Demonstrator Analysis

In order to find out about the specific characteristics of each security risk assessment task, the corresponding process from the IEC 62443-3-2 standard has been used for three practical security risk assessments of typical industrial demonstrators inside the SmartFactoryOWL[1], which is a research and technology transfer factory operated by Fraunhofer IOSB-INA and OWL University of Applied Sciences and Arts, in Lemgo, Germany. Each of the corresponding security risk assessment tasks has varying characteristics, such as financial cost, time, knowledge and experience required, relation to stakeholders, or the dependence towards other tasks. Therefore, the conducted practical security risk assessments were used to determine the status quo with regard to a possible automation of certain tasks of the process model.

The first inspected demonstrator (see Figure 5.1) is a developed and implemented system from the research project DEVEKOS[2], which inherits a novel skill-based engineering and communication scheme based on OPC UA with real-time capabilities and vendor-independent functionalities. The system contains around 30 different controllers from six manufacturers and produces fidget spinners providing a real-world example of an industrial production system. According to [191], it can be classified as a learning factory and following the ISA 88 classification, it is a process cell.

The second investigation was performed on an industry-grade towel folding machine (see Figure 5.2) produced by the company Kannegiesser provided via the ADIMA[3] research project. This machine contains only one controller and the communication architecture is provided by one vendor resulting in a much simpler system. It can be

[1] https://smartfactory-owl.de/?lang=en

[2] https://www.devekos.org/projektdemonstrator

[3] https://www.init-owl.de/en/research/projects/detail/adaptives-assistenzsystem-fuer-die-instandhaltung-intelligenter-maschinen-und-anlagen

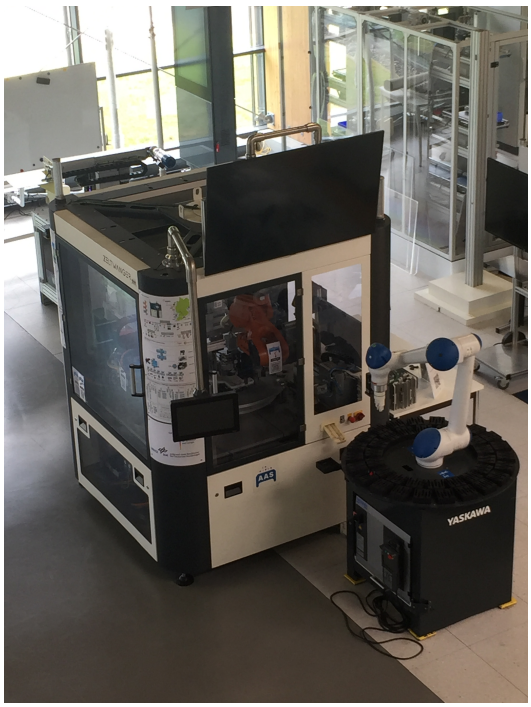

Figure 5.1: OPC UA demonstrator in the SmartFactoryOWL from the DEVEKOS research project

classified as an industrial development extent [191] and as a unit following the ISA 88 definition.

The third security risk assessment was performed on the Customisable Production System demonstrator (see Figure 5.3) within the AutoS^2 research project[4]. A subpart of the demonstrator was inspected including a laser engraving cell for LEGO figurines, an external exhaustion unit, a conveyor belt for product transportation, and the associated cabinet hosting the control, networking, infrastructure, and communication hardware. This system consists of one process cell according to ISA 88 with one external interface to the Internet and includes three controllers from one common manufacturer. It can also be classified as a learning factory [191].

The following categorisation named Level of Knowledge (LOK) is defined to be able to evaluate the needed expert knowledge and the difficulty of each security risk assessment task. It is adapted from the IEC 62443-3-3 standard and the corresponding SL specifications of typical attacker motivations and resource usage ranging from zero to four. Hence, the LOK definition is based on agreed and accepted qualitative metrics from within the security domain. Following the standard, SL 0 is chosen when there are no security measures at all. SL 1 delivers protection against casual or coincidental violation, SL 2 provides protection against intentional violation using

[4]https://www.init-owl.de/en/research/projects/detail/automatische-bewertung-und-ueberwachung-von-safety-security-eigenschaften-fuer-intelligente-technische-systeme

Figure 5.2: Industry-grade towel folding machine in the SmartFactoryOWL from the ADIMA research project

simple means, SL 3 gives protection against intentional violation using sophisticated means, and finally SL 4 is described as protection against intentional violation using sophisticated means with extended resources [135].

Every LOK ranging from zero to four can also be described by an average financial cost depending on the required skills and resources estimated from typical loans inside the consulting domain because most of the asset owners do not employ security experts and are dependant on external analysts. The LOK values for the respective ZCRs are defined based on the experiences from the practical security risk assessments and out of the discussions with domain experts and typical asset owners. The calculated financial costs (based on needed time and required LOK) are used to make the assessments comparable.

- Junior Analyst (LOK 0): No specific requirements or security-related skills → 80€/hour
- Analyst (LOK 1): Simple means with minimum resources and basic security-related skills → 100€/hour
- Senior Analyst / Junior Security Analyst (LOK 2): Simple means with low resources and generic security-related skills → 120€/hour

- Security Analyst (LOK 3): Sophisticated means with moderate resources and OT specific skills → 150€/hour
- Senior Security Analyst (LOK 4): Sophisticated means with extended resources and OT specific skills → 200€/hour

Figure 5.3: Customisable Production System demonstrator in the SmartFactoryOWL from the AutoS^2 research project

5.1.2. Case Study Summary

Each security risk assessment task (represented by a ZCR from the IEC 62443-3-2) was performed on all three demonstrating systems (1) DEVEKOS, (2) ADIMA, and (3) AutoS^2 accordingly in a fully manual way. Table 5.1 shows the overall summary of the results acquired during this process. The ZCRs are listed with the corresponding tool coverage and the linked LOA values as originally defined in Section 4.1: Documentation = 1 / Checklists = 2 / SIEM = 3 / Passive OT monitoring = 4 / Active scanners = 5 / Research and development = 6. In addition, the experienced LOK values are presented and used to calculate the financial cost for each practical security risk assessment task based on the measured time required for the certain ZCR. The ZCR 2.1 is be used here as an explanatory example requiring 13.0h during the security risk assessment of the first system from the DEVEKOS project. This task can be further described with a LOA of 3 and a LOK of 4 representing a task for a typical senior security analyst with OT specific skills and extended resources assisted by available software tools, e.g. passive OT monitoring. The total financial cost for this ZCR at the inspected system are 2.600€, calculated with the measured 13.0h multiplied with 200€ per hour as a typical loan in Germany based on the LOK.

Table 5.1: Practical evaluation results from the SmartFactoryOWL for the time and financial cost requirements of the (1) DEVEKOS, (2) ADIMA, and (3) AutoS^2 research projects

RM Task	Tool Category	LOA	LOK	(1) Time [h]	(1) Cost [€]	(2) Time [h]	(2) Cost [€]	(3) Time [h]	(3) Cost [€]
ZCR 1.1	3 & 5	1	2	6.0	720	1.0	120	4.5	540
ZCR 2.1	2, 4 & 6	3	4	13.0	2600	1.5	300	6.0	1200
ZCR 3.1	6	0	4	5.0	1000	1.5	300	2.0	400
ZCR 3.2	6	2	4	1.0	200	0.5	100	0.5	100
ZCR 3.3	6	2	4	1.0	200	0.5	100	0.5	100
ZCR 3.4	6	2	4	1.0	200	0.5	100	0.5	100
ZCR 3.5	6	2	4	1.0	200	0.5	100	0.5	100
ZCR 3.6	6	2	4	1.0	200	0.5	100	0.5	100
ZCR 4.1	-	2	3	2.0	300	2.0	300	1.0	150
ZCR 5.1	6	0	4	8.0	1600	1.5	300	2.0	400
ZCR 5.2	4, 5 & 6	3	3	10.0	1500	2.0	300	3.0	450
ZCR 5.3	6	0	4	2.0	400	0.5	100	1.0	200
ZCR 5.4	6	0	4	2.0	400	0.5	100	1.0	200
ZCR 5.5	6	0	4	1.5	300	0.5	100	0.5	100
ZCR 5.6	-	0	3	5.0	750	1.0	150	3.0	450
ZCR 5.7	-	0	3	1.0	150	0.5	75	0.5	75
ZCR 5.8	6	0	4	2.0	400	1.0	200	2.0	400
ZCR 5.9	-	0	4	3.0	600	0.5	100	1.0	200
ZCR 5.10	-	0	4	2.0	400	0.5	100	1.0	200
ZCR 5.11	-	0	4	1.5	300	0.5	100	0.5	100
ZCR 5.12	-	0	4	3.0	600	1.0	200	2.0	400
ZCR 5.13	1, 4 & 6	3	2	2.0	240	0.5	60	1.0	120
ZCR 6.1	1	0	3	1.0	150	0.5	75	1.0	150
ZCR 6.2	1	1	2	0.5	60	0.5	60	0.5	60
ZCR 6.3	1	0	4	0.5	100	0.5	100	0.5	100
ZCR 6.4	1	0	3	1.0	150	0.5	75	1.0	150
ZCR 6.5	1	0	3	0.5	75	0.5	75	0.5	75
ZCR 6.6	1	0	3	1.0	150	0.5	75	1.0	150
ZCR 6.7	1	0	3	0.5	75	0.5	75	0.5	75
ZCR 6.8	1	0	3	0.5	75	0.5	75	0.5	75
ZCR 6.9	1 & 2	1	3	0.5	75	0.5	75	0.5	75
ZCR 7.1	1, 2 & 6	1	4	5.0	1000	1.0	200	4.0	800
Total amount:				**85.0**	**15170**	**24.5**	**4290**	**44.5**	**7795**

The further analysis of Table 5.1 reveals six ZCRs which are of main interest based on the combination of a low LOA due to the absence of available tools, a need on domain-specific know-how represented by a high LOK, and great resource requirements with regard to time and cost. The following list summarises the dominant six identified ZCRs together with their respective justification:

1. ZCR 2.1 "Perform initial cyber security risk assessment"
   - Medium LOA available
   - High LOK necessary
   - High resource requirements
2. ZCR 3.1 "Establish zones and conduits"
   - Low LOA available
   - High LOK necessary
   - Medium resource requirements
3. ZCR 5.1 "Identify threats"
   - Low LOA available
   - High LOK necessary
   - High resource requirements
4. ZCR 5.2 "Identify vulnerabilities"
   - Medium LOA available
   - Medium LOK necessary
   - High resource requirements
5. ZCR 5.6 "Determine SL-T"
   - Low LOA available
   - Medium LOK necessary
   - Medium resource requirements
6. ZCR 7.1 "Attain asset owner approval"
   - Medium LOA available
   - High LOK necessary
   - Medium resource requirements

In order to analyse the results in a more detailed way, Figure 5.4 shows the mean percentages of the identified ZCRs with regard to the needed time for each security risk assessment task performed at the three different systems from the SmartFactoryOWL. The focused six ZCRs represent 18.75% of the whole ZCR

process (32 ZCRs are present within the IEC 62443-3-2 in total) but make up for nearly half (44.6%) of the needed time within the conducted security risk assessments. Figure 5.5 presents a similar picture based on the mean percentages with regard to financial costs. The six identified ZCRs make up for 46.6% of the related financial costs. This biased tendency supports the statement of having a few ZCRs functioning as bottlenecks which should be the focus for the upcoming research to automate certain tasks for industrial security risk assessments.

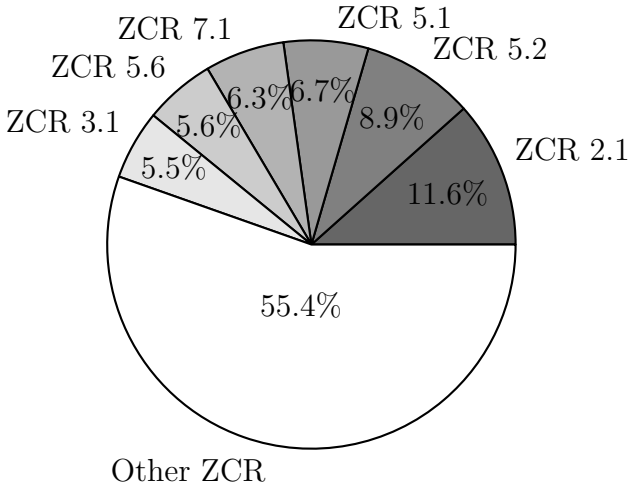

Figure 5.4: Mean percentages of the three security risk assessments with regard to the time requirements of the ZCRs

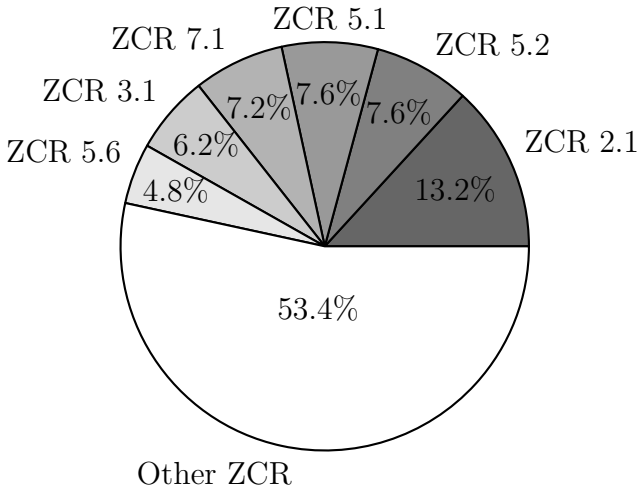

Figure 5.5: Mean percentages of the three security risk assessments with regard to the financial cost requirements of the ZCRs

In addition, the ZCR 2.1 should function as an explanatory example. The purpose of the initial cyber security risk assessment is to understand the worst-case scenarios

present to the SUC of the organisation. Currently, for industrial assets there are only tools available for the identification of technical vulnerabilities to support this task. The LOA equals 3 with regard to the tool category of passive OT monitoring and the needed LOK is set to 4 based on the task description and the corresponding requirements for analysts performing this typical security risk assessment task. The measured time is the highest of the practical evaluation. These aspects combined indicate many potentials for huge gains with regard to the automation of security risk assessments.

5.2. Focus for Automation

The results of the three security risk assessment case studies described in the section before are used for the identification of required efforts regarding time and financial cost, needed resources, required know-how, and possible bottlenecks during the overall process. The various ZCRs and the associated activities represent high and diverse demands towards security experts [44]. Therefore, this dissertation focuses on the selected ZCRs to achieve improvements towards the automation of security risk assessments where the expected results will have the highest impact. The previous case studies (also refer to [53]) revealed that only six of the total 32 specific ZCRs are consuming up to nearly 50% of the required time and financial cost:

- ZCR 2.1 "Perform initial cyber security risk assessment"
- ZCR 3.1 "Establish zones and conduits"
- ZCR 5.1 "Identify threats"
- ZCR 5.2 "Identify vulnerabilities"
- ZCR 5.6 "Determine SL-T"
- ZCR 7.1 "Attain asset owner approval"

Hence, these six identified ZCRs are the starting point regarding the method for information and process modelling towards the automation security risk assessments. In general, the various tasks for security risk assessments result in a high demand for resources, such as personnel, finances, and comprehensive knowledge, so that a focus on certain tasks is essential to achieve first results [44]. Figure 5.6 summarises the overall scope and the alignment of the six identified ZCRs into a four-phased process in conformity to the IEC 62443-3-2 standard.

The overall method for information and process modelling (collection, formalisation, usage, and access) proposed in this dissertation, is performed on the six identified and focused ZCRs and described within the following chapters 6 to 9. This represents an exemplary utilisation of the four method steps based on the most resource-intensive tasks of the security risk assessment process that have the most rewarding possibilities

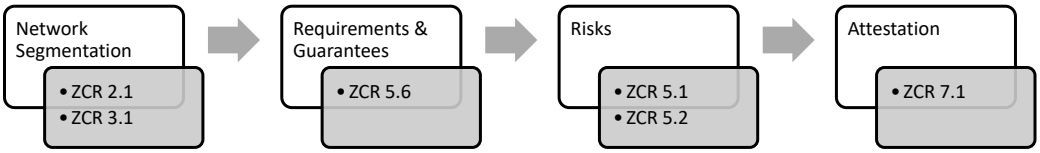

Figure 5.6: Process for the initial security risk assessment enhanced with five additional ZCRs from the IEC 62443-3-2 standard

for automation and subsequently reducing manual efforts. In general, the IEC 62443-3-2 standard defines two abstraction levels of security risk assessments: (1) initial and (2) detailed. This dissertation covers the initial abstraction level (comparable to the Initial Cybersecurity Risk Assessment (ICRA) from the ISA-TR 84.00.09), specified within the ZCR 2.1, enhanced with the additional five ZCRs from the detailed abstraction level. The overall purpose is to gain an initial understanding of the unmitigated risks of the SUC [73]. Furthermore, the remaining ZCRs not in scope of the evaluated method are only analysed in a theoretical way and need to be added in future work to accomplish a complete automation of the IEC 62443-3-2 standard. Nevertheless, no contradictions or drawbacks of the proposed method based on the following concepts were experienced and therefore indicate a general usability towards the remaining ZCRs:

- Information Collection (see Chapter 6): Swimlanes
- Information Formalisation (see Chapter 7): UML class diagrams
- Information Usage (see Chapter 8): Predicate logic
- Information Access (see Chapter 9): AAS

The security risk assessment process automated consists of four phases (Network Segmentation / Requirements & Guarantees / Risks / Attestation) integrating, adapting, and fulfilling the six identified ZCRs (2.1 / 3.1 / 5.1 / 5.2 / 5.6 / 7.1) of the IEC 62443-3-2 standard. By mapping the ZCRs towards more abstract phases, the intrinsically linked characteristics of certain ZCRs can be regarded, e.g. the combined performance of ZCR 5.1 "Identify threats" and ZCR 5.2 "Identify vulnerabilities", similar to the risk identification proposal of a Detailed Cybersecurity Risk Assessment (DCRA) from the ISA-TR 84.00.09. The focus on threats as a basis for security risk assessments is also supported by the work of [2]. Nevertheless, the order of the ZCRs was slightly adjusted in comparison to the IEC 62443-3-2 standard to emphasise the focus on the ZCR 5.6 "Determine SL-T" utilisation due to the negligence of this metric in practice so far and in contrast the high potential which is present to express stakeholder demands identified within the ongoing research project AutoS^2. A similar change of the ZCR scope is also done within the ISA-TR 84.00.09

document regarding the DCRA definition. By doing so, the mapping towards the IEC 62443-3-2 standard is still ensured and a higher degree of practicability can be achieved due to the experiences from the security risk assessment case studies mentioned before in Section 5.1.

This results in a functional and usable level of completeness, although the overall achievements provided by this dissertation do not replace an exhaustive manual security risk assessment in presence, including e.g. penetration testing on-site or requested workshops with the asset owner. Nevertheless, the automation represents a draft regarding the security state of the SUC as a support basis for further decision making by the respective stakeholders. This results in an automation of a limited scope and a dedicated focus of tasks regarding an initial security risk assessment enhanced with additions towards a detailed one (see IEC 62443-3-2 definitions). The method is able to perform first-time security risk assessments and also re-assessments in a manner typically described in the consulting business as a hybrid version of a high-level (initial following the IEC 62443-3-2 definition) and a low-level (detailed following the IEC 62443-3-2 definition) security risk assessment. Afterwards, the associated asset owner or operator of the SUC can then use the documented attestation of the automated security risk assessment to integrate it into the corresponding RM processes to decide about follow-up risk treatment, e.g. accepting, transferring, mitigating, or terminating the risks (as defined within the ISO/IEC 27005 standard). The described automated process uses safety as a security objective based on the associated asset characteristics. An alignment with typical safety processes is not yet regarded, but is available as a draft within the ISA-TR84.00.09-2023 document.

The IEC 62443-3-2 standard can be referred to as an open guideline [74] and describes the inherited tasks (ZCRs) and the corresponding information only on a generic level. The necessary details for a specific implementation which is required for automation are left out on purpose. Therefore, a further analysis of the required inputs, involved stakeholders, order of decisions, and generated outputs needs to be done for each ZCR. Security expert knowledge that is defined for one ZCR is also regarded as given in all subsequent ZCRs [44]. This results in an expert system for automated security risk assessments including the six identified ZCRs in focus.

6. Information Collection

The following chapter covers the first method step of information collection as the basis towards automated security risk assessments. First, in Section 6.1 a theoretical analysis is performed to identify and document the requirements regarding the security risk assessment process. Then, Section 6.2 shows the used swimlane methodology as an adapted version of the Business Process Model Notation (BPMN) to collect all necessary information based on the four phases of network segmentation (Section 6.2.2), requirements & guarantees (Section 6.2.3), risks (Section 6.2.4), and attestation (Section 6.2.5). Finally, Section 6.3 concludes the presented contents by providing an intermediate summary of the achievements.

6.1. Requirements Analysis: Information Collection

This section contains the fact sheets for requirements resulting from the related work analysis and the literature research regarding the automation of security risk assessments, representing the method step of information collection (abbreviated with "IC"). Each requirement (abbreviated with "R") can be identified via its short name as the title in bold and a unique identifier in the following style "R-IC-#". Furthermore, a description of the requirement is given to summarise the demanded contents. The three different verbs which are used, represent the relevance of each requirement in a prioritising order (in accordance to the RFC 2119[1] and the ISO/IEC Directives[2]): High = *"shall"* / Medium = *"should"* / Low = *"may"*. All requirements are not directly measurable and are therefore evaluated in a qualitative manner based on a subjective justification by the author of this dissertation later on. Finally, associated and generally covered statements from other references are listed to further support the requirements analysis.

Requirement: Security Conformity

- ID: R-IC-1
- Description: The security risk assessment *shall* conform to the state of the art in OT security.
- Generally covered statements from other references:

[1] https://datatracker.ietf.org/doc/html/rfc2119
[2] https://www.iso.org/sites/directives/current/part2/index.xhtml

  - The security risk assessment shall align to the approaches from the IEC 62443 standard [57].
  - The use of common metrics for evaluation eases the combination with other approaches [192].
  - An information model shall be interoperable and not proprietary [44].
  - Attack modelling is required for security risk assessments [192].
  - Security risk analyses shall be present to address essential gaps in a cost-effective manner [192].
  - Result consistency shall be achieved by using objective and generally-accepted approaches [154].
  - Fixing security knowledge of a certain point in time into an implementation is doomed to fail [193].
  - Security risks shall be evaluated based on assets, vulnerabilities, and threats [54].

Requirement: Decision Making

- ID: R-IC-2
- Description: The security risk assessment result *should* be adequate to support human decision making.
- Generally covered statements from other references:
  - Typical target groups have limited security knowledge and only limited amount of time for security-related decisions [28].
  - The results should always be available for the involved decision makers [154].
  - The communication of results should increase the awareness of stakeholders [65].
  - The results should include humans for the final decision making [28].
  - The security risk assessment should utilise a functional approach integrating the human decision makers [183].
  - Quantitative security risk analyses are objective, dependent on probabilities, confusing, complex, time consuming, and require more preliminary work [110].

Requirement: Use Cases

- ID: R-IC-3
- Description: The security risk assessment *should* be applicable for various use cases regarding different SUCs.

- Generally covered statements from other references:
  - The concepts should be adaptive towards the upcoming frequent changes of network architectures [154].
  - Maintaining security during operation to gain efficiency and to be adaptable to the ever-changing security landscape [57, 183].
  - The target type of the security risk assessment should be industrial networks [154].
  - The goal of the security risk assessment should be compliance [154].
  - The approach should function in a dynamic and adaptable manner [154].
  - A higher usefulness and acceptance can be expected if various use cases are covered [192].
  - The expert knowledge should be described generically [44].

Requirement: Safety Coupling

- ID: R-IC-4
- Description: The security risk assessment *should* be coupled with the safety domain and the associated information.
- Generally covered statements from other references:
  - The connection and interdependence of safety and security has to be addressed [192].
  - Coupling of safety and security risk assessments is demanded to acquire future-proof approaches [31].

Requirement: Resource Consumption

- ID: R-IC-5
- Description: The security risk assessment *may* be resource-efficient.
- Generally covered statements from other references:
  - The approach may be efficient and cost-effective to increase the overall applicability and usefulness [192].
  - The computational requirements may be cost-effective. [154].
  - The complexity of security risk assessment phases and their expected duration may be decreased [192].

Requirement: Data Quality

- ID: R-IC-6
- Description: The security risk assessment *may* be based on high-quality data.

- Generally covered statements from other references:
  - Use up-to-date security expert knowledge [44].
  - The required data may be easily accessible and ideally system-immanent [194].
  - Machine-readable data formats and modelling languages [154].
  - Independence of platforms, software, or OSs [154].
  - Parallel understandability for humans and machines [44].
  - The data may represent security-relevant information from IACSs [28].
  - The security knowledge may be correct and provided in time [44].

6.2. Swimlane Specification

This section builds upon previous works by the author of this dissertation and extends the published results from [56, 57, 94].

6.2.1. From Process Modelling to Swimlanes

This subsection aims to describe the first method step regarding the collection of information which are required to translate the abstract level of the IEC 62443-3-2 standard into a practical and tangible semi-formal information model for an automated security risk assessment. The result will be a collected information base fulfilling the demands of the IEC 62443-3-2 standard and providing the fundamentals for further formalisation, usage, and access of the information respectively. In summary, the essential question is which specific information is necessary to perform a security risk assessment in compliance to the IEC 62443-3-2 process description in scope of this dissertation to be prepared for automation.

To develop the abstract concepts from the IEC 62443-3-2 standard up to the degree necessary for automation, an analysis followed by a specification needs to be performed. This work uses a lightweight and adapted version of the BPMN typically used by business analysts as a tool to collect and visualise all information mandatory for the investigated process. By doing so, a visual overview for easy understanding of the IEC 62433-3-2 security risk assessment process can be provided. BPMN typically works from a process-oriented view fitting well to the needed perspective, providing an as-is model that reflects the current state of contents. During this method step of information collection, the goal is to get a better understanding and a detailed overview of the security risk assessment contents to fulfil the demands from the IEC 62443-3-2 standard. Therefore, the BPMN approach to specify the associated swimlanes can be seen as a tool to progress with the proposed method steps and not as part of the information model. Nevertheless, UML activity diagrams would also have been applicable here, but they require a more strict modelling language resulting in a slower and more tedious specification of the information models. The

typical definition process of a BPMN model consists of the following parts in the given order further shown in Figure 6.1.

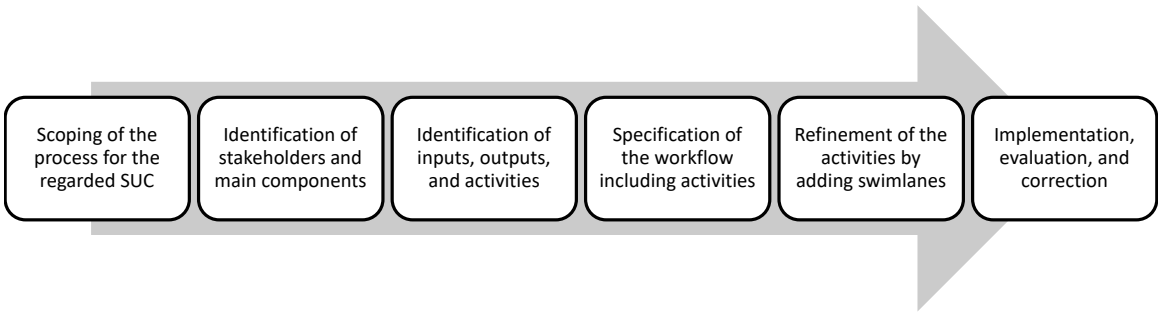

Figure 6.1: Definition process to create BPMN models based on swimlanes [56]

The result is a representation within swimlanes (similar to a UML activity diagram) containing an overview of the stakeholders, inputs, decisions, outputs, information model elements, and environmental influences in a graphical way. The main structure is represented by a pool containing the different process stakeholders. In this case, there are six swimlanes inside the pool. Each swimlane includes one of the three process stakeholders i.e., (1) component manufacturer, (2) system integrator, and (3) asset owner, as well as (4) the information model representing the internal data for processing, (5) the security risk assessment process as the basis for the decision logic, and (6) the environment containing external data, e.g. from public databases or already formalised security expert knowledge. In general, the swimlanes can exchange information with each other using the arrow connectivity object as an indication for the sequence flow representing the order and the connections between all objects within the pool. Other general objects within the pool are described and used as follows [94]:

- Start/End: The start object defines the beginning and the end object defines the result of the whole process. Both must occur at least once per process. This object is represented by the rectangular with evened edges.
- Activity: An activity is the most common of all objects within the pool. It describes the plethora of tasks and the entirety of activities is linked to a chain with the sequence flows. This object is represented by a rectangular with sharp edges.
- Decision: These objects are used to define decisions within the pool in the form of a "gateway" with defined inputs, outputs, and internal logic. They are reliant on inputs and typically have two outputs following a boolean (true or false) decision regarding the question stated within the gateway. This object is represented by the rhomboid.
- Input/Output: The inputs and outputs of the pool are associated with the corresponding swimlanes i.e., process stakeholders. They contain the name

and the type of the variable. These objects are represented by the coloured rectangulars.

Figure 6.2 shows an excerpt from the complete definition of the swimlanes as an exemplary guidance of the overall process. The presented swimlane part includes the abstracted and shortened example of the ZCR 5.6 "Determine SL-T" from the requirements and guarantees phase. It includes the six presented swimlanes representing the important stakeholders, the inputs and outputs, the start of the complete process, the related activities, and the decisions to be taken [94]. The detailed description of the example can be found within the Annex B in Section 15. This can be seen as a blueprint for swimlane definition process and the related tasks. The detailed figures of the swimlanes for each of the security risk assessment phases can be found within the Annex A in Section 14.

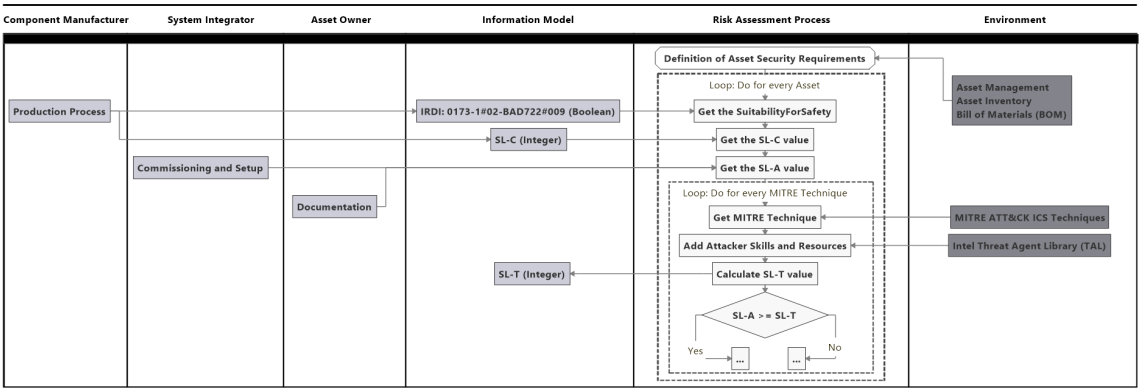

Figure 6.2: Abstracted example of the general swimlane definition [94]

By using the swimlanes in the presented way, also an alignment towards the defined requirements from Section 6.1 can be achieved. The final evaluation of the requirements analyses can be found within Chapter 11.

- Security Conformity (R-IC-1): The specification of the security risk assessment process within the swimlanes includes all necessary tasks from scoping to documentation and the results are traceable. In addition, only widely-accepted and generally-used metrics are integrated to achieve a high interoperability and credibility. Furthermore, the specification maps to the contents and requirements from the IEC 62443-3-2 standard and the ZCRs in scope.
- Decision Making (R-IC-2): The documentation of the security risk assessment results is done via an attestation specifically targeted at typical OT stakeholders to give them insights into their SUC and to support their security-related decision making. Therefore, a qualitative approach was chosen to create objective results that are not biased by expert opinions and creating results in a timely and cost-efficient manner.

- Use Cases (R-IC-3): The security risk assessment process is able to cover various use cases if the preconditions of information modelling and data availability are provided. The whole approach can be used during different lifecycle phases of typical industrial systems from engineering, commissioning, or operation in a resource-effective way. By doing so, several use cases can be tackled to achieve a compliance towards security requirements.

- Safety Coupling (R-IC-4): The specified process within the swimlanes highly focuses on the security objective of safety by using safety-relevant asset characteristics to determine the target assets in scope. This enables a light coupling between the safety and security domain. Nevertheless, the presented concept does not cover any safety risk assessment process or calculation of safety metrics, such as SILs or PLs.

- Resource Consumption (R-IC-5): By eliciting security expert knowledge into a software-based security risk assessment process, a higher efficiency and a reduction of required resources can be achieved. In addition, typical operators from the OT domain are enabled to take part in security-related decision making and the general security risk assessment processes for their systems. Furthermore, the computational requirements of the automated process are very low and can also be deployed in resource-restricted environments.

- Data Quality (R-IC-6): The security risk assessment process is based on DTs with live data from the SUC to guarantee an up-to-date and correct information base. These system-immanent information are extended and enhanced with external and publicly available frameworks. By intertwining the usage of machine-readable data formats and human-readable result documentation, a hybrid and parallel comprehensibility is given.

6.2.2. Phase 1: Network Segmentation

The first phase of network segmentation (ZCR 2.1 & ZCR 3.1) includes the initial analysis of the SUC including all assets, their topology, and their characteristics with safety as the main security objective (as described within Section 3.4). The aim is to include existing zones (logical grouping of assets with similar security requirements, e.g. safety relevance) and conduits (logical definition of interfaces between zones) or to define new ones as a basis for further assessment. In addition, assets need to be attributed towards the responsible and accountable personnel. The following list shows the most important facts about the associated ZCRs:

- ZCR 2.1 Requirement: Perform initial cyber security risk assessment of the SUC to identify the worst-case unmitigated risks

- ZCR 2.1 Rationale
  - Initial understanding of compromising possibilities

  - Prioritisation of detailed security risk assessments
  - Facilitation of zone and conduit definition
  - Integration of safety-relevant analyses
- ZCR 2.1 Information
  - Asset characteristics, especially safety-related
  - Typical operation procedures
  - Modular manufacturing system architecture
  - Interfaces and communication capabilities
- ZCR 3.1 Requirement: Establishment of zones and conduits
- ZCR 3.1 Rationale
  - Identification of assets with common security requirements
  - Effective evaluation and implementation of mitigation measures
- ZCR 3.1 Information
  - Physical and logical asset topology
  - Responsible employee
  - Accountable employee

This phase is started with the SUC definition as a preparation based on the necessary basic information provided including a Bill of Material (BOM) representing the machine, modules, and assets in scope of the security risk assessment (stored within the associated AASs). Figure 6.3 shows an abstract overview of the first phase of network segmentation including all involved activities as defined within the swimlanes.

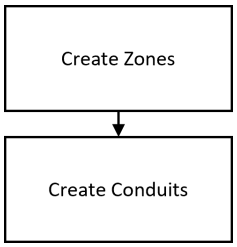

Figure 6.3: Abstract overview of the swimlanes for the first phase of network segmentation

The first main activity contains the creation of zones on an intra-modular level based on the asset characteristic of the suitability for safety functions. The module represents the scope of the production process that the associated asset is involved

in and used for. This information typically comes from the asset owner who has the insights into the details of the production process. In addition, the asset characteristic of the suitability for safety functions is used to specify the zones regarding security. Currently, the focus is set on Ethernet-based communication paths due to the prevalent presence within networking architectures in scope (as described within Section 3.2 [10]). Assets without such an interface, e.g. serial, proprietary, or bus-based protocols, are not integrated into the security risk assessment process. Still, an adaptation towards wireless communication is generally possible. This focus can be further explained with the example of a bus coupler with Ethernet communication functionality, which is typically an asset without safety characteristics. Nevertheless, if the bus coupler is connected via a bus to analogue or digital inputs or outputs being responsible for safety-critical tasks within the SUC, the bus coupler is treated as an asset relevant for safety aggregating the characteristics of the connected devices. By doing so, up to now, two zones per module are created, one for safety-related assets and one for non-safety-related assets. To be conformant to the IEC 62443-3-2, every zone needs to have an identification number, an accountable employee, and a responsible employee assigned.

The second main activity is the creation of conduits. Based on the physical connections of assets derived from the information within the AAS, the communication interfaces between zones are identified. The assets on the edges of the zones are categorised as access point assets for the later security risk assessment process and added to the conduit definitions. Furthermore, every conduit has an identification number, an accountable, and a responsible employee assigned to be compliant with the IEC 62443-3-2.

6.2.3. Phase 2: Requirements & Guarantees

The second phase (ZCR 5.6) includes the specification of applicable security requirements, such as possible attack vectors, and guarantees, such as implemented mitigation measures. This results in a security-oriented view on the desired target level (SL-T) for each module to match it later on to the currently achieved state of security [195]. Similar approaches are known from other domains as well, e.g. a privacy classification of information regarding the GDPR or a general criticality classification for business processes and organisations. The SL-T determination is brought forward and prioritised here compared to the original IEC 62443-3-2 process to emphasise the integration of the SL-T for the subsequent security risk assessment tasks. The following list shows the most important facts about the associated ZCR:

- ZCR 5.6 Requirement: SL-T determination for each zone or conduit
- ZCR 5.6 Rationale
  - SL-T as the desired target state of security
  - Communication method towards responsible stakeholders

  - Expression is possible as a single value or as different styles of vectors

- ZCR 5.6 Information
  - Possible attackers including their characteristics and capabilities
  - Possible threats

This phase of the security risk assessment consists of four main activities starting with the initialisation of the required knowledge base containing all necessary information for the evaluation of requirements and guarantees, e.g. MITRE ICS ATT&CK techniques and mitigations plus their mapping towards the IEC 62443 FRs, more specifically towards the CRs and SRs. In addition, the utilised IEC 62443 vectors are created and prepared for the security risk assessment. Figure 6.4 shows an overview of the second phase of requirements and guarantees.

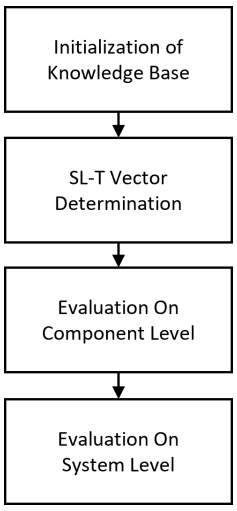

Figure 6.4: Abstract overview of the swimlanes for the second phase of requirements and guarantees

The second main activity covers the determination of the SL-T vector (as proposed within the IEC 62443-3-2 standard) containing all single SL-T values as a preparation for the security risk assessment. This dissertation uses the SL-T vector as a basis for the definition of security requirements for comparison with the security guarantees later on. The two remaining main activities can be grouped together based on their contents. To achieve a complete conformity with the IEC 62443 standard, two other SL vectors are required: SL-C and SL-A. The required information need to be provided by the component manufacturer for the SL-C during the manufacturing process and by the system integrator for the SL-A during the commissioning and implementation processes on the site of the asset owner. These information are assumed to be available and need to be present for a complete security risk assessment. Afterwards, the evaluation is performed in two parts. First, the evaluation is

performed on the component level and afterwards on the system level. All CRs of every identified asset of the SUC are initialised with a start value. Then, the capable (C) and achieved (A) SL vectors are compared to decide which CRs are mitigated on the component level and which ones need to be shifted to be evaluated on the system level. In addition, it can be checked whether there are any assets present which already achieve the desired SL-T, but can be reconfigured to reach a higher SL. Subsequently, the evaluation of the zones takes place on the system level. All unmitigated CRs from the component level are then investigated again. Therefore, the access point assets as the entry points to the zones are evaluated if they can fulfil the so far unmitigated requirements. Three results are possible here: (1) The SR will remain with an unmitigated status, (2) the SR can be further improved by a reconfirmation, or (3) the SR is marked as mitigated and does not need to be checked again during the following security risk assessment tasks.

6.2.4. Phase 3: Risks

The third phase (ZCR 5.1 & ZCR 5.2) describes the identification, analysis, and evaluation of security risks which result from technical vulnerabilities that can be exploited by threat actors with different skills, resources and motivations. The output is a list of unmitigated risks including information on impact and complexity. The characteristic of complexity is used here in favour of the likelihood because there is a huge uncertainty regarding available probabilities of human attackers and no reliable data is existing [196]. In addition, existing results from previous security risk assessments are planned to be included as a basis for new security risk assessments as well. The following list shows the most important facts about the associated ZCRs:

- ZCR 5.1 Requirement: Identification of threats potentially affecting zones or conduits
- ZCR 5.1 Rationale
  - Comprehensive and realistic list of threats
  - Grouping and further classification is advised due to large amount of possible threats
  - Identification of potentially affected assets, zones, or conduits
- ZCR 5.1 Information
  - Threat source
  - Capabilities or skill-level of the threat source
  - Possible threat vectors
- ZCR 5.2 Requirement: Identification and documentation of vulnerabilities
- ZCR 5.2 Rationale

  - Vulnerabilities are required to better understand the threat vectors
  - The usage of public databases and reports is advised
  - Also manual vulnerability assessments can be performed
- ZCR 5.2 Information
  - Vulnerability characteristics

Figure 6.5 shows an overview of the third phase regarding the security risks. Essentially, this phase consists of seven main activities collecting, analysing, and determining the risks of all involved asset types, namely access points, path assets, and targets. The whole security risk assessment is based on the present network topology and the available asset characteristics. In addition, external information, such as vulnerabilities and threat actor data, are used to enhance the system-related information.

The seven main activities can be split up into three categories, each covering all of the different asset types. The first category of main activities includes the collection of all access point, target, and path assets from the complete SUC. Therefore, the complete BOM is scanned for assets with the previously assigned access point characteristic, i.e. assets on the edge of a zone and at the two ends of a conduit. These assets are especially interesting because they are functioning as an entry spot towards each zone from a security point of view. In the next task, all target assets are identified based on their suitability for safety functions. These target assets are the central objective of the defined security risk assessment process. Lastly, the path assets are identified based on the prevalent network topology and the physical connection of the assets. These connections are determined by the physical ports of the assets and information on these ports is derived from the respective AASs and the associated submodels.

The second category of main activities covers the identification and analysis of technical vulnerabilities of access point, path, and target assets. Currently, all external information needed for this security risk assessment task are prepared in the form of a local knowledge base, including CVE and CVSS attack vector and scope metrics for vulnerabilities, MITRE ICS ATT&CK techniques and mitigations, and associated Intel TAL skills and resources. All of these required information are open-source and publicly available. Therefore, they can be retrieved from Internet-based databases, such as the NIST National Vulnerability Database (NVD)[3] or the German equivalent VDE CERT[4].

The last category only contains one main activity that is responsible for the determination of security risks associated with the previously identified and analysed assets. All target assets are covered and inspected by specifying an attack path through the network based on the access point assets as entries to the zone and the path assets as intermediate connections towards the final target asset. Currently,

[3]https://nvd.nist.gov
[4]https://cert.vde.com/en

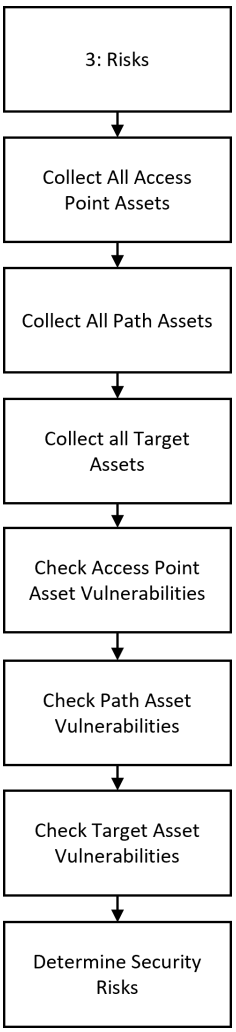

Figure 6.5: Abstract overview of the swimlanes for the third phase regarding risks

the prototypical implementation only covers line network topologies, but the general method and concept proposal works for other topologies, such as ring, mesh, or star, as well. All possible paths based on present technical vulnerabilities and unmitigated techniques towards target assets are compared and each target asset is enhanced with a CVSS impact and a CVSS attack complexity. This results in the final security risk determination based on typical risk matrices, e.g. as defined within the informative Annex B of the IEC 62443-3-2 standard.

6.2.5. Phase 4: Attestation

The fourth and last phase is the attestation (ZCR 7.1) which documents the results in the form of findings with the aim of attaining the final approval by the responsible and accountable stakeholders and the agreement of tasks following the security risk

assessment [37]. In addition, an adequate and understandable attestation is needed to achieve credibility for the results from the automated process. The following list shows the most important facts about the associated ZCR:

- ZCR 7.1 Requirement: Attainment of the asset owner approval
- ZCR 7.1 Rationale
  - Presentation of the security risk assessment results
  - Review of the results by the asset owner
  - Decisions about further risk actions
- ZCR 7.1 Information
  - Organisational facts for documentation

This phase includes only one main activity to collect all the necessary information required for the attestation and is therefore not further supported with a figure. Generally, two different kinds of information are collected. Firstly, the general data of the security risk assessment is included, such as date and time, identification number, errors that occurred, computing time, and the current operator. Secondly, the security-relevant data are included, such as the total number of modules, components, zones, targets, and assessed CRs respectively SRs, the highest security risk of all target assets, the lowest Intel TAL attacker skill and resource category, and the resulting risks for all targets. More details can be found within Section 10.2.2 regarding the prototypical implementation and the documentation of the attestation.

6.3. Intermediate Summary

By completing the first method step of information collection as described in the previous sections, the basis for improving security risk assessments towards automation is available. This addresses the deficit D1 of insufficient process coverage from Section 4.3.3. The definition of the swimlanes includes the specification of the abstract process from the IEC 62443-3-2 standard and the elicitation of security expert knowledge into a description which is no longer reliant on human input. Due to a focus on the contents within the scope for this dissertation based on three practical analyses in the SmartFactoryOWL and a theoretical analysis of the associated standards, six out of 32 ZCRs in total are chosen to be used for the information and process modelling method. Through the usage of widely accepted security concepts, such as IEC 62443-3-2 SLs, CVE and CVSS metrics, MITRE ATT&CK ICS techniques and mitigations, or the Intel TAL, a high credibility and transparency can be achieved when no proprietary information models are developed. Figure 6.6 shows the overall dissertation structure again and the current progress regarding the first method step of information collection.

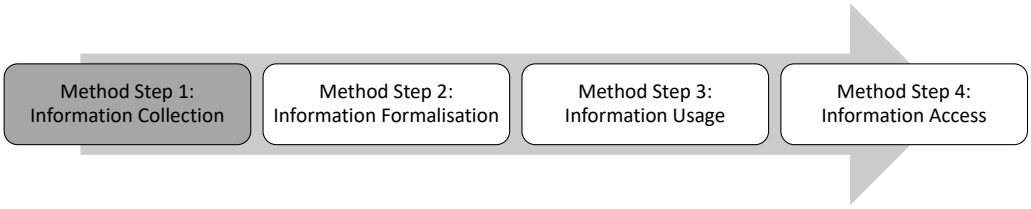

Figure 6.6: Dissertation progress after the first method step of information collection

The method step of information collection covers the general knowledge engineering steps of task identification and collection of relevant knowledge [58]. To perform the remaining method steps towards the automation of security risk assessments, machine-readable information is required. Therefore, the next method step within the upcoming chapter describes the formalisation of the collected information.

7. Information Formalisation

The overall goal of this second method step is the technology-independent formalisation of the collected information into a semi-formal model for the automation of security risk assessment processes defined within the IEC 62443-3-2 [57, 197]. This formalisation of informal security expert knowledge is needed to reduce required efforts and necessary resources by means of automation [44, 154, 160]. In general, models always represent an investigated system within some domain of discourse up to a certain degree of detail and correctness, point of view, and purpose, in this case the security risk assessment of modular IACSs [198]. By modelling this process, a semantic description of security expert knowledge is possible [44] which helps to remove biased decision-making processes [110], increases confidence and completeness of security risk assessments [114], and enables efficient flexibility and interoperability based on self-description [6]. This is described within the following chapter, covering the second method step of information formalisation.

First, in Section 7.1 a theoretical analysis is performed to identify and document the requirements regarding the information model for this dissertation. Section 7.2 shows the defined UML class diagrams based on the swimlane specification to formalise the information into a semi-formal model based on the four phases of network segmentation (Section 7.2.1), requirements & guarantees (Section 7.2.2), risks (Section 7.2.3), and attestation (Section 7.2.4). Finally, Section 7.3 concludes the presented contents by providing an intermediate summary of the achievements and next steps.

7.1. Requirements Analysis: Information Formalisation

This section contains the fact sheets for requirements resulting from the related work analysis and the literature research regarding the automation of security risk assessments, representing the method step of information formalisation (abbreviated with "IF"). Each requirement (abbreviated with "R") can be identified via its short name as the title in bold and a unique identifier in the following style "R-IF-#". Furthermore, a description of the requirement is given to summarise the demanded contents. The three different verbs which are used, represent the relevance of each requirement in a prioritising order (in accordance to the RFC 2119[1] and the ISO/IEC Directives[2]): High = *"shall"* / Medium = *"should"* / Low = *"may"*. All requirements are not directly measurable and are therefore evaluated in a qualitative manner

[1] https://datatracker.ietf.org/doc/html/rfc2119

[2] https://www.iso.org/sites/directives/current/part2/index.xhtml

based on a subjective justification by the author of this dissertation later on. Finally, associated and generally covered statements from other references are listed to further support the requirements analysis.

Requirement: Credibility

- ID: R-IF-1
- Description: The information model *shall* have a high credibility based on completeness and correctness.
- Generally covered statements from other references:
  - Security decisions shall be possible to be tracked and to reconstruct how and why they have been made [28].
  - Security risk analyses shall incorporate a repeatable process which results in the same risk determination regardless of who performs the analysis [2].
  - Completeness of the security risk assessment and correctness of the results shall be guaranteed [114].
  - Unbiased decision making shall be ensured [58].
  - A model shall have clearly specified scales consisting of the type of used scales, the range of values, and their respective interpretation [154].
  - A continuously checked and validated internal state is essential for the correct process of decision making [154].
  - Security models should reflect all possible internal system states [154].
  - The information model shall be independent of an implementation [44].

Requirement: Knowledge Representation

- ID: R-IF-2
- Description: The information model *shall* support the rule-based formalisation and elicitation of expert knowledge.
- Generally covered statements from other references:
  - Knowledge-based expert systems are widely distributed and used to represent knowledge and rules [44].
  - A clear separation of activities, domains, tools, and interfaces between stakeholders is necessary [74].
  - The model shall be able to document the process results [44].

Requirement: Reusability

- ID: R-IF-3

- Description: The information model *should* be developed in a compositional manner.
- Generally covered statements from other references:
  - Usage of pre-defined libraries as a digital representation [183].
  - As many elements as possible of the information model as the decision base should be re-usable [28].
  - Modularity and adaptability are important for possible future updates [44].
  - The information model should be developed in a compositional approach to allow for replacement of various parts of the model [114].
  - Hierarchisation should help to reach a complexity reduction [44].
  - Interchangeability and reusability of security risk assessments should be ensured [74].

Requirement: Training Data

- ID: R-IF-4
- Description: The information model *should* not require training data.
- Generally covered statements from other references:
  - There is a lack of empirical research and high-quality historical data usable as background knowledge [62].
  - The majority of available data is either based on opinions, anecdotal evidences, or experiences [110].
  - The assessment of ICS security risks is currently limited by the available data [65].
  - Training data will never be holistic and the residual risks s_{n+1} cannot be regarded [65].

7.2. UML Class Diagrams

Out of the presented formalisms from Section 4.2, the syntactic and semantic usage of UML is chosen for the information formalisation and modelling. Currently, UML is the standard notation for various types of stakeholders, such as developers, project managers, business owners, or technical experts, to model complex characteristics in an easy and human-readable manner [198]. The current UML version is 2.5 and comprises 14 different diagrams in total, offering a plethora of characteristics and possibilities which can be categorised into two types, namely (1) structure diagrams and (2) behaviour diagrams [59]. Structure diagrams (also called static semantics) mainly describe the organised objects of the described model at a specific

point in time, whereas behaviour diagrams (also called dynamic semantics) are emphasising on the dynamics of a given model and its changes over time [198]. In this dissertation, class diagrams from the structural type are utilised to model the associated information regarding the security risk assessments due to their fit towards the requirements defined within Section 7.1.

The usage of UML class diagrams has the following main advantages. This modelling approach enables a detailed specification of the required information structure, helps to understand complex interrelations, increases visibility, and enhances the quality in a human-readable and visualised way [59, 168, 173, 198]. In addition, UML is used in various domains independent of businesses or industries and can be used for nearly all phases of specification, development, and implementation [168]. In an ideal case, the usage of UML increases the understandability of the described model and reduces time and efforts for comprehension and realisation by providing a common ground for information exchange even between available software tools on the market [59, 174]. In general, the UML specification is documented in an extensive way, but to start and understand UML only a fraction of the contents is required and an introduction to the modelling approach with UML is very intuitive for humans [173].

For completeness, its disadvantages need to be mentioned here as well. In some cases a formal notation using UML is not really necessary to specify an easy problem [59]. Also, the UML specification document is very large and experts discuss whether all the presented information are really required and whether they are comprehensible. [198]. In addition, there are no formulas or hints for a recommended amount of UML diagrams or the inherited objects which leads to various degrees of detail and extensive diagrams tend to become more difficult to understand for first-time readers. Finally, UML has the same drawback regarding actuality as all other specifications resulting in additional efforts to maintain the created diagrams to be up-to-date [59]. Nevertheless, the described disadvantages do not impact the modelling results in this work.

The specification of the UML class diagrams is done via a manual analysis and translation of the swimlanes from Section 6.2 based on the provided inputs, outputs, variable types, and the association between the various objects. The required information for the security risk assessment are analysed and sorted for the creation of distinctive objects always in alignment with the IEC 62443-3-2 concepts, e.g. *Zone*, *Conduit*, *Risk*, or *Asset*. Afterwards, these objects are translated into classes within the UML diagrams and the associated information are turned into UML properties for the corresponding classes. Properties can have among others one of the following five primitive data types: Integer, Boolean, String, UnlimitedNatural, or Real [198]. This enables the usage of the classes from the UML information models to match the needed contents from the four phases of the security risk assessment process, namely network segmentation, requirements & guarantees, risks, and attestation. Certain important classes, such as *Zone*, *Conduit*, or *Asset*, are used in multiple phases of the complete process. Therefore, they appear in multiple UML class diagrams of the

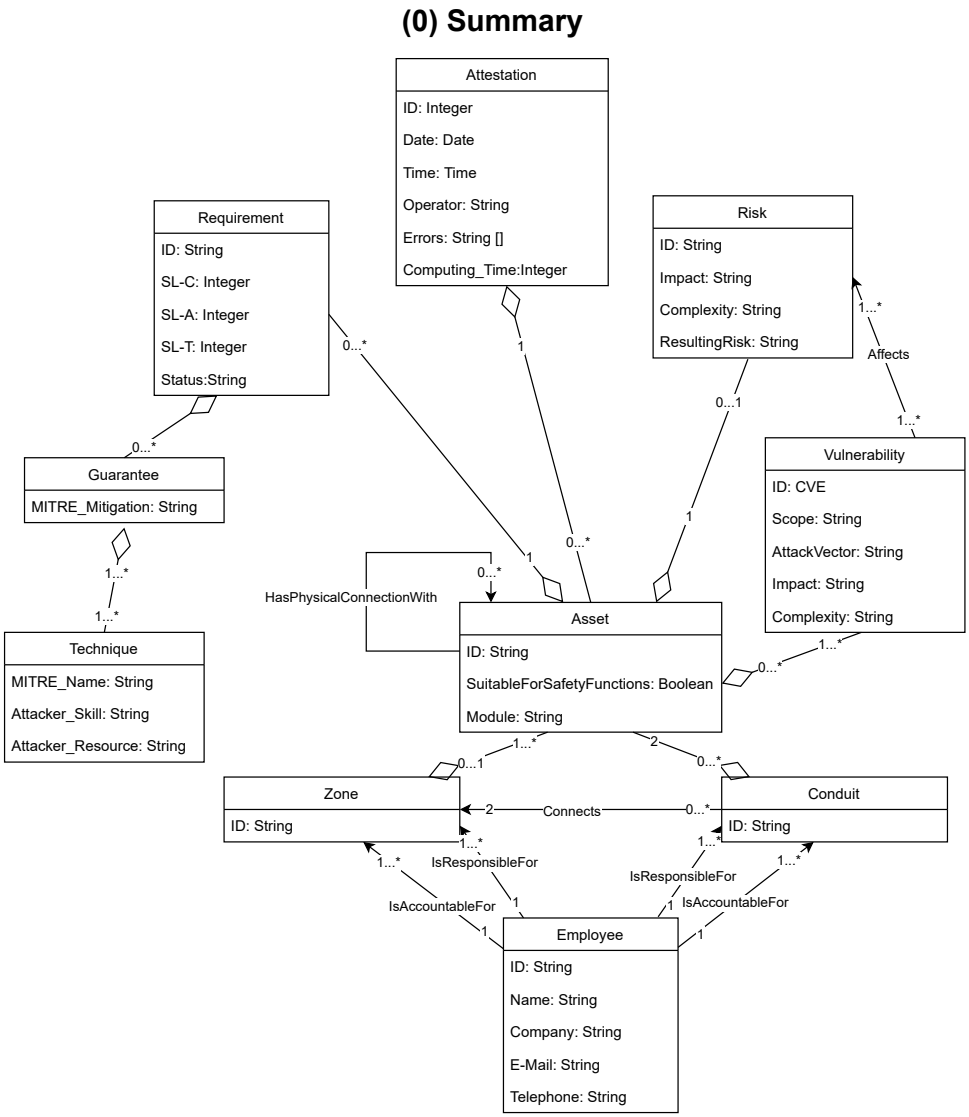

Figure 7.1: UML class diagram summarising the information model for all four phases

phases with identical properties. By doing so, also an alignment towards the defined requirements from Section 7.1 can be achieved:

- Credibility (R-IF-1): Modelling with the UML notation offers the possibility to check for completeness in a human-readable manner resulting in a appropriate credibility due to reproducible and traceable characteristics.
- Knowledge Representation (R-IF-2): Using UML enables the formalisation, modelling, and elicitation of security expert knowledge in a specified way to make it available for subsequent utilisation for automating security risk assessment processes.

- Reusability (R-IF-3): The individual UML class diagrams developed for each of the four phases result in a compositional structure which can be inspected separately for each phase or as a summarised view on all phases together. In addition, the single diagrams can be edited or exchanged without impacting other parts of the information model in an abstracted and hierarchical way.
- Training Data (R-IF-4): The manual approach of creating the UML class diagrams as the main information model does not require any training data which is an advantage as no training data is available at this stage. In addition, this enables a white box behaviour.

Figure 7.1 shows the final UML class diagram combined for all four phases of the modelled security risk assessment process comprising the network segmentation, requirements & guarantees, risks, and attestation. During the combination, duplications from the various phases are removed and the overall figure is created. In general, each UML class diagram consists of objects and their relationships. The objects are certain classes necessary for the information model and generated out of the swimlanes, whereas the relationships follow a clear definition from the UML specification. In total, two types of relationships are used: (1) association (`->`) reflecting a general and directed connection between classes needing a reference among them to describe a common goal and (2) aggregation (`-<>`) as a subset of association implying more details for the relation between classes being child and parent, but being able to exist independently of each other. They are further described by the cardinality at the start or the end of the arrow and the optional label, e.g. "Conduit 0...* `->` 2 Zone" defines that each conduit connects two zones and every zone can have zero to infinite conduits. Each of the following subsections shows the contents for a specific security risk assessment phase highlighted in grey in a chronological manner based on a repetitive structure explaining the modelled objects, their characteristics, and their interrelations.

7.2.1. Phase 1: Network Segmentation

This subsection builds upon previous works by the author of this dissertation and extends the published results from [56, 57, 94]. The following Figure 7.2 shows the formalisation of the collected information for the first phase of network segmentation as a UML class diagram.

Fact sheet for the first phase of network segmentation [57]:

- Reference: Section 6.2.2 containing the swimlane specification
- Four classes
  1. *Asset* (Three properties)
  2. *Zone* (One property)
  3. *Conduit* (One property)

4. *Employee* (Five properties)

- Used UML relationships
  1. Association (->)
  2. Aggregation (-<>)

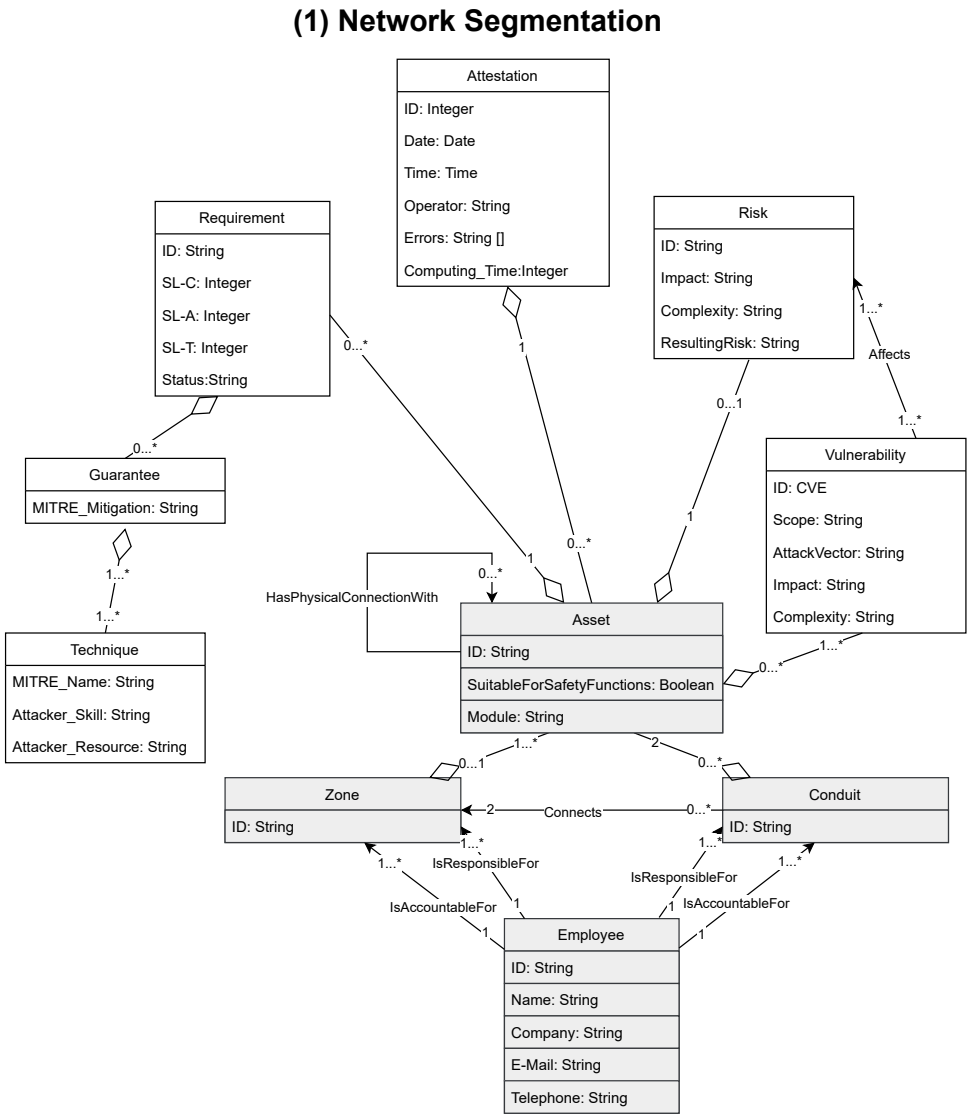

Figure 7.2: UML class diagram for the first phase of network segmentation

The four specified classes (*Zone*, *Conduit*, *Employee*, and *Asset*) with their corresponding properties cover all information needed for the first phase of network segmentation. The aim here is the analysis of the network topology and the definition of zones and conduits based on the asset characteristic safety represented as an ECLASS property. An asset is understood here as a typical industrial hardware

component belonging to a dedicated machine module, such as a sensor, actuator, or controller. The coupling of security and safety is further discussed within Section 3.4 and is already covered in other publications, such as [31] or [199]. ECLASS is an intersectoral classification system based on international standards describing products, materials, or services and their properties with unique identifiers and a well-fitting example of an information description format to be reused to achieve a high credibility and interchangeability [200]. The property "SuitableForSafetyFunctions" of the *Asset* class can be used to cover the demanded information about an asset's safety capabilities. The boolean characteristic (True / False) is described in the ECLASS standard (International Registration Data Identifier (IRDI) 0173-1#02-BAD722#009) and assigned to the asset by the component manufacturer. The focus of the example is set here on the safety property, but also other characteristics can be represented by ECLASS properties, e.g. the asset ID by the IRDI 0173-1#02-ABA684#002 for unique identification of an asset in the context of its overall application. In addition, a property describing the assignment of an *Asset* to a module (String) is utilised for the network segmentation into zones and conduits. Furthermore, each *Asset* gets a unique identifier (ID). The same holds true for each *Zone*, *Conduit*, and *Employee* class receiving an ID. Furthermore, the *Employee* is described by additional String properties (name, company, e-mail, and telephone) following the IEC 62443-3-2 standard.

7.2.2. Phase 2: Requirements & Guarantees

Fact sheet for the second phase of requirements & guarantees:

- Reference: Section 6.2.3 containing the swimlane specification
- Six classes
  1. *Technique* (Three properties)
  2. *Mitigation* (One property)
  3. *Requirement* (Five properties)
  4. *Asset* (One property)
  5. *Zone* (One property)
  6. *Conduit* (One property)
- Used UML relationships
  1. Association (`->`)
  2. Aggregation (`-<>`)

Figure 7.3 shows the formalisation of the collected information for the second phase of requirements & guarantee as a UML class diagram:

The classes *Asset*, *Zone*, and *Conduit* are used and integrated into the UML information model as specified within the previous subsection about the first phase

Figure 7.3: UML class diagram for the second phase of requirements & guarantees

of network segmentation. In addition to that, three new classes are required to fulfil the needs for this phase of the security risk assessment and the formalisation of requirements & guarantees: *Technique*, *Mitigation*, and *Requirement*. The main idea of this phase (see Section 6.2.3) is to compare possible *Techniques* which might be used to attack the SUC with the associated *Mitigations* which are used to protect the SUC. Each *Technique* and *Mitigation* is enhanced with a unique name (String) which is based on the naming scheme and conventions from the MITRE ATT&CK ICS framework. The UML classes here were also created in alignment with the Annex D of [44] specifying the threat and protection measure classes. In addition, the *Technique* class uses the previously introduced Intel TAL to inherit the

specific attacker skills and resources. The class of *Requirements* is used to collect the three different types of SLs from the IEC 62443-3-2 standard (SL-C = capable, SL-A = achieved, and SL-T = target) plus an ID containing either CR or SR plus the corresponding number and a status property. The usage of these classes is further explained within [94]. By using these approved security concepts instead of proprietary ones, a gain in credibility and transparency can be achieved and the sharing of expert knowledge becomes easier, resulting in an increased security overall. Furthermore, the specification of pre-defined knowledge bases or libraries, preferably as a digital representation, will improve the reusability and integration into other approaches [183].

7.2.3. Phase 3: Risks

Fact sheet for the third phase regarding risks:

- Reference: Section 6.2.4 containing the swimlane specification
- Eight classes
  1. *Vulnerability* (Five properties)
  2. *Risk* (Four properties)
  3. *Requirement* (Five properties)
  4. *Mitigation* (One property)
  5. *Technique* (Three properties)
  6. *Asset* (Three properties)
  7. *Zone* (One property)
  8. *Conduit* (One property)
- Used UML relationships
  1. Association (`->`)
  2. Aggregation (`-<>`)

Figure 7.4 shows the formalisation of the collected information for the third phase of risks as a UML class diagram:

The classes *Asset*, *Zone*, *Conduit*, *Technique*, *Mitigation*, and *Requirement* are used and integrated into the UML information model as specified within the previous subsections. In addition to that, two new classes are required to fulfil the needs for this phase of the security risk assessment and the formalisation of risks: *Vulnerability* and *Risk*. The goal of this phase (see Section 6.2.4) is the identification, analysis, and evaluation of risks based on the results from the previous subsections. Therefore, the *Vulnerability* class is formalised in alignment with CVSS and added to the UML class diagram. It includes a unique ID for identification and four other properties as String enumerations adapted from the CVSS base metric group: scope, attack

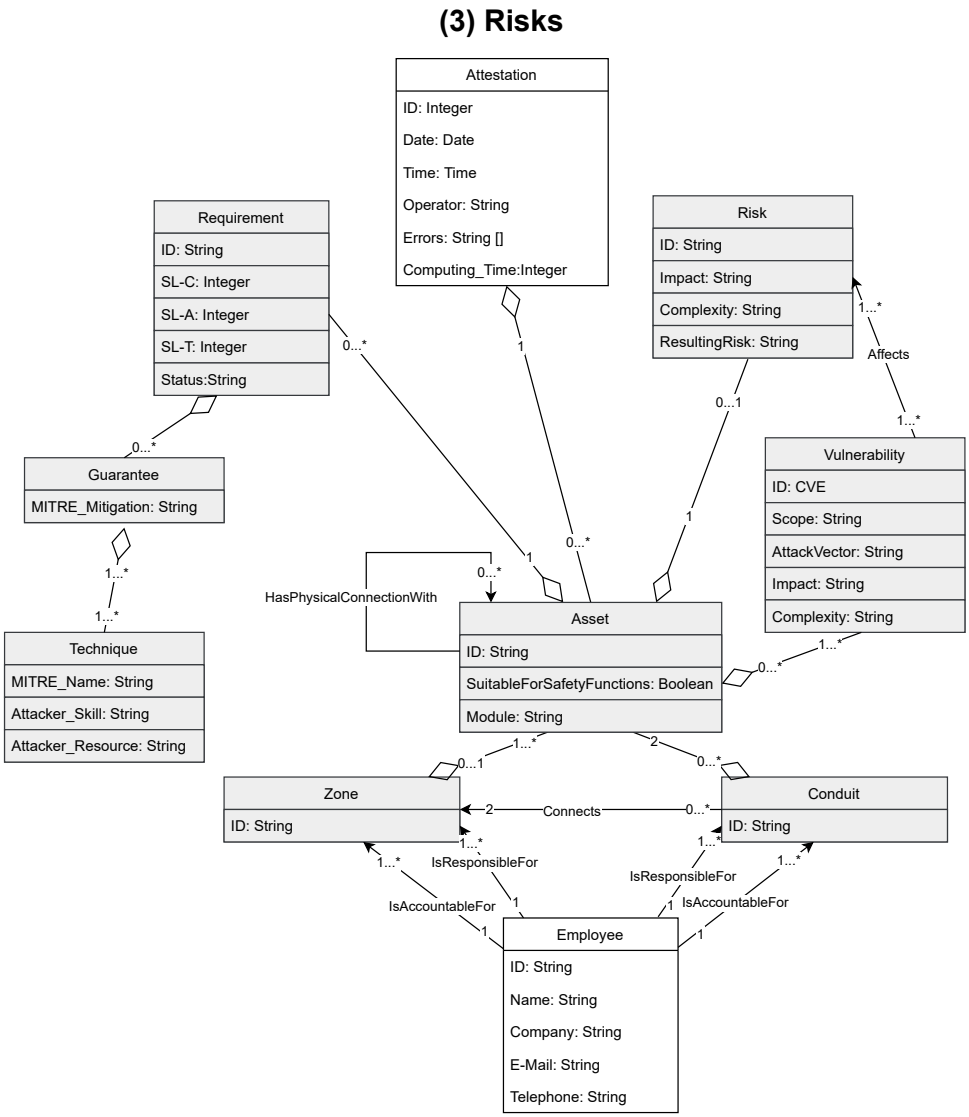

Figure 7.4: UML class diagram for the third phase regarding risks

vector, impact, and complexity. These properties describe typical characteristics of a *Vulnerability* and they are used to assess the effects of a *Vulnerability* onto the SUC and the associated security risk assessment. The following explanatory statements are excerpts adapted from the original CVSS specification[3]:

- Scope: This metric captures whether a vulnerability in one vulnerable component impacts resources in components beyond its security scope. Formally, a security authority is a mechanism (e.g., an application, an OS, firmware, a sandbox environment) that defines and enforces access control in terms of

[3]https://www.first.org/cvss/v3.1/specification-document

how certain subjects/actors (e.g., human users, processes) can access certain restricted objects/resources (e.g., files, CPU, memory) in a controlled manner. All the subjects and objects under the jurisdiction of a single security authority are considered to be under one security scope. If a vulnerability in a vulnerable component can affect a component which is in a different security scope than the vulnerable component, a scope change occurs. Intuitively, whenever the impact of a vulnerability breaches a security/trust boundary and impacts components outside the security scope in which vulnerable component resides, a scope change occurs. Possible metric values are: Unchanged (U) and Changed (C).

- Attack Vector: This metric reflects the context by which vulnerability exploitation is possible. This metric value (and consequently the base score) will be larger the more remote (logically, and physically) an attacker can be in order to exploit the vulnerable component. The assumption is that the number of potential attackers for a vulnerability that could be exploited from across a network is larger than the number of potential attackers that could exploit a vulnerability requiring physical access to a device. Possible metric values are: Network (N), Adjacent (A), Local (L), and Physical (P).
- Impact: This metric captures the effects of a successfully exploited vulnerability on the component that suffers the worst outcome that is most directly and predictably associated with the attack. Analysts should constrain impacts to a reasonable, final outcome which they are confident an attacker is able to achieve. Only the increase in access, privileges gained, or other negative outcome as a result of successful exploitation should be considered when scoring the impact metrics of a vulnerability. Possible metric values are: High (H), Low(L), and None (N).
- Complexity: This metric describes the conditions beyond the attacker's control that must exist in order to exploit the vulnerability, e.g. the collection of more information about the target, or computational exceptions. Importantly, the assessment of this metric excludes any requirements for user interaction in order to exploit the vulnerability. Possible metric values are: High (H) and Low (L).

After evaluating the present *Vulnerabilities* within the inspected *Assets* of the SUC, *Risks* are identified, analysed, and evaluated as the main concept of the security risk assessment process. The *Risk* class contains four properties in total. First, a unique ID is specified. Afterwards, the three necessary properties of impact, complexity, and the resulting risk are filled. The impact and complexity directly depend on the CVSS metrics mentioned before. The resulting risk is calculated via a typical risk matrix in alignment to the Annex B of the IEC 62443-3-2 standard, which is shown in Table 7.1. The combination of the impact and complexity properties result in a qualitative risk evaluation based on a five-tier ranking: Very low, low, medium, high, and very high.

Table 7.1: Risk matrix based on CVSS characteristics used for the security risk assessment in alignment with Annex B of the IEC 62443-3-2 standard

Complexity / **Impact**	**Low**	**High**
None (N)	Low	Very Low
Low (L)	Medium	Medium
High (H)	Very High	High

7.2.4. Phase 4: Attestation

Fact sheet for the fourth phase of attestation:

- Reference: Section 6.2.5 containing the swimlane specification
- Nine classes
  1. *Attestation* (Six properties)
  2. *Risk* (Four properties)
  3. *Requirement* (Five properties)
  4. *Mitigation* (One property)
  5. *Technique* (Three properties)
  6. *Asset* (Three properties)
  7. *Zone* (One property)
  8. *Conduit* (One property)
  9. *Employee* (Five properties)
- Used UML relationships
  1. Association (`->`)
  2. Aggregation (`-<>`)

Figure 7.5 shows the formalisation of the collected information for the fourth phase of attestation as a UML class diagram:

The classes *Asset*, *Zone*, *Conduit*, *Technique*, *Mitigation*, *Requirement*, and *Risk* are used and integrated into the UML information model as specified within the previous subsections. In addition to that, one new class is required to fulfil the needs for this phase of the security risk assessment and the necessary documentation: *Attestation*. The overall aim of this phase (see Section 6.2.5) is to collect and document all relevant information from the previous three phases regarding the results of the security risk assessment. The term *Attestation* is derived and chosen in accordance with the documentation patients are receiving when visiting a doctor to get an update about their health status. Furthermore, the term certification is highly biased and already used for several other security-related concepts. All in all,

Figure 7.5: UML class diagram for the fourth phase of attestation

this process step is necessary to gain credibility with the accountable and responsible stakeholders. In addition, follow-up actions, such as the definition of risk treatment options or the improvement of the inspected SUC, can be based on the *Attestation*. The *Attestation* class contains six properties in total. First, a unique ID is specified. Then, the remaining five properties are integrated to document the meta data of the software-based automation of the specific security risk assessment execution: Date, time, operator, errors, and the computing time.

7.3. Intermediate Summary

The second method step of information formalisation further enhances the collected information base towards the automation of security risk assessment processes. The specified swimlanes from the previous chapter are taken as the foundation for the information modelling based on UML class diagrams and in accordance with already established and well-acknowledged formalisms from the security domain, including the definition of objects, properties, and their relationships. This addresses the deficit D2 of missing standardised metrics from Section 4.3.3. Therefore, the first step is to analyse and document the requirements towards the desired information model. These requirements are then considered for the choice of the type of modelling formalism used, here in favour of UML class diagrams. Afterwards, all four phases of network segmentation, requirements & guarantees, risks, and attestation are modelled in distinctive UML class diagrams according to the respective swimlane contents. In addition, a summarising UML class diagram is presented to cover all the modelled information in one common figure. All of the class diagrams include the corresponding classes with the associated properties to describe their characteristics and the relationships between them. Figure 7.6 shows the overall dissertation structure again and the current progress regarding the second method step of information formalisation.

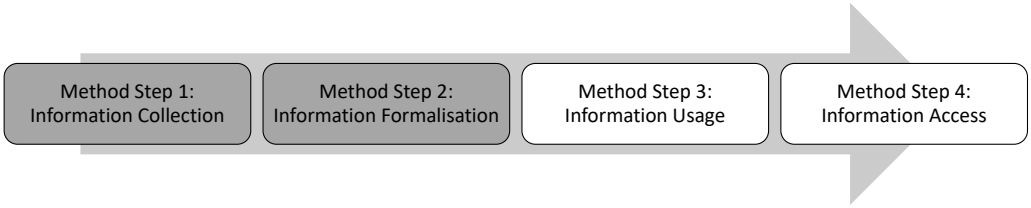

Figure 7.6: Dissertation progress after the second method step of information formalisation

This method step of information formalisation aligns with the general knowledge engineering process step of deciding on a fitting vocabulary from [58]. To perform the remaining method steps towards the automation of security risk assessments, the elicitation of expert knowledge needs to be continued in alignment with the creation of machine-readable logic as a basis for decision-making. Therefore, the next method step described within the upcoming chapter describes the logic-based usage of the collected and formalised information.

8. Information Usage

The following chapter covers the third method step of information usage as an intermediate step towards automated security risk assessments. First, the associated requirements for this method step are specified within Section 8.1 based on available references and literature. Sections 8.2.1 and 8.2.2 motivate the usage of logic-based rules and compare the available logic languages to decide on a fitting solution. The analysis includes normal applications, case-based systems, rule-based systems, AI, and ML respectively [44]. Afterwards, Section 8.2.3 presents the characteristics and the summary for the rule specification. Furthermore, each of the four security risk assessment phases of network segmentation (Section 8.2.4), requirements & guarantees (Section 8.2.5), risks (Section 8.2.6), and attestation (Section 8.2.7) is explained and illustrated. Finally, Section 8.3 concludes the presented contents by providing an intermediate summary of the achievements and next steps.

8.1. Requirements Analysis: Information Usage

This section contains the fact sheets for requirements resulting from the related work analysis and the literature research regarding the automation of security risk assessments, representing the method step of information usage (abbreviated with "IU"). Each requirement (abbreviated with "R") can be identified via its short name as the title in bold and a unique identifier in the following style "R-IU-#". Furthermore, a description of the requirement is given to summarise the demanded contents. The three different verbs which are used, represent the relevance of each requirement in a prioritising order (in accordance to the RFC 2119[1] and the ISO/IEC Directives[2]): High = *"shall"* / Medium = *"should"* / Low = *"may"*. All requirements are not directly measurable and are therefore evaluated in a qualitative manner based on a subjective justification by the author of this dissertation later on. Finally, associated and generally covered statements from other references are listed to further support the requirements analysis.

Requirement: Explainability

- ID: R-IU-1
- Description: The rule set as a basis of the security risk assessment results *shall* be explainable, interpretable, and trustworthy.

[1] https://datatracker.ietf.org/doc/html/rfc2119
[2] https://www.iso.org/sites/directives/current/part2/index.xhtml

- Generally covered statements from other references:
  - The rule set needs to offer the possibility to inspect it and to understand why a certain result was achieved based on the provided input [58].
  - The definition and usage of the rules shall be explainable [63].
  - Implementations need to have methods to prevent unsafe states, e.g. version control, (unit) testing, reviews, or monitoring [58].
  - The achieved results of the rule set shall be accountable [58].
  - It needs to be clear which rules and information are created by human experts and which ones are based on learning approaches [66].
  - The objectives of the security risk assessment shall be clear and fixed towards safety, a tradeoff due to varying objective functions is not allowed [58].

Requirement: Expert Knowledge

- ID: R-IU-2
- Description: The logic-based rules *shall* capture various forms of expert knowledge and heuristics in an intuitive way.
- Generally covered statements from other references:
  - Heuristics are an effective way for decision-making when the full facts are not known [201].
  - The expert knowledge needs to be collected in a comprehensible way [44].
  - Logic-based approaches can be used to capture human reasoning [193].
  - Explicit knowledge is available in an abstract form and can be passed on independently to others, e.g. a security guideline or standard [44].
  - Implicit knowledge typically cannot be described by exact words [44].
  - The model shall support forward-chained rules [44].

Requirement: Rule Characteristics

- ID: R-IU-3
- Description: The knowledge elicitation *should* be easy to use and unambiguous.
- Generally covered statements from other references:
  - Knowledge representation languages should be declarative, compositional, expressive, context independent, and unambiguous [58].
  - The input of knowledge into the model needs to be simple and the development effort should be low [44].
  - The usage of logic inherits a low complexity, shorter development cycles, and less computational resources [58].

8.2. Logic-Based Knowledge Elicitation

The specification of the logic-based rules is done via a manual analysis and translation of the decisions within the swimlanes from Section 6.2 based on the provided inputs, outputs, and variable types. By doing so, it is possible to capture and elicit the security expert knowledge necessary for the security risk assessment steps in a logic-based rule format. This enables an easy specification requiring low efforts and a format which is simultaneously human-readable and machine-readable. In addition, the rules are used to specify the software components of the prototypical implementation and the required decisions that are further explained in Section 10.2. This enables the usage of logic-based rules to match the needed contents from the four phases of the security risk assessment process, namely network segmentation, requirements & guarantees, risks, and attestation. This section builds upon previous works by the author of this dissertation and extends the published results from [56, 57].

8.2.1. Motivation for Logic Usage

In general, expert knowledge can be captured in various ways, but it is typically stored as facts, rules, object descriptions, heuristics, or conditions [60]. The analysis within the work [44] shows a clear preference of rule-based expert systems for security analyses over commonly used normal applications, e.g. Microsoft Excel from the Office suite[3] or case-based systems. This preference is due to two reasons: (1) A separation between knowledge and programming logic and (2) a lack of predefined cases and missing transferability between use cases. Additionally, the use of approaches from the domain of AI is infeasible as there are currently no public datasets available which describe the processes of security risk assessments [196, 202]. Furthermore, the present requirements from the safety domain prohibit a usage of black box systems where it is not known how certain results were achieved due to missing transparency [61]. The lack of labelled datasets at this stage prevents the use of self-learning algorithms for the present problem statement. Creating suitable labelled datasets is also not within the scope of this dissertation, therefore the use of AI is not further considered here. In addition, the results of the automated security risk assessments need to have a high level of traceability, explainability, and credibility which cannot be provided by self-learning algorithms so far. This leads into the research domain of logical languages to address the problem of security risk assessments and the rule-based utilisation of the previously collected and formalised information.

In general, a large part of the required expert knowledge for security risk assessments is available in a rule-like (if-then-relationship) format or can be transferred to such a format [46]. In addition, the use of a rule-based system offers the advantage that the knowledge base can be flexibly and quickly adapted regarding the high degree of possible changes, e.g. newly discovered technical vulnerabilities or utilised

[3] https://www.microsoft.com/de-de/microsoft-365/excel

attack vectors by human adversaries [46]. Also the work by [54] shows that algebraic specifications are favoured for security-related notions. This results in updated security risk assessments always being carried out with the latest state of knowledge in each process step [46]. By doing so, the amount of required knowledge remains manageable, the needed time for analyses is reduced, the management of complex systems becomes possible, and up-to-date security approaches can be integrated [46]. Therefore, the stated requirements from Section 8.1 are fulfilled as described here:

- Explainability (R-IU-1): The utilisation of a rule-based approach based on a fixed set of rules defined by human experts offers a high trustworthiness [58]. In addition, the results of the security risk assessment are always repeatable, understandable, explainable, and accountable [63, 66].
- Expert Knowledge (R-IU-2): The logic-based rule definition uses a widely disseminated and established semantic to represent knowledge in a way which is similar to human reasoning [44]. It is possible to specify the information model based on explicit and implicit expert knowledge depending on the data source.
- Rule Characteristics (R-IU-3): The deduction of rules from the information model enables a declarative and expressive way of representing the knowledge [58]. The required resources to create the rules and the overall development efforts are low [44, 58].

There are various research works available which integrate and use rules for security in general. One of the first works that uses a logic- and rule-based approach is [203] from the year 2002, where security statements are modelled with regard to authorisation via, e.g. certificates, policies, or Access Control Lists (ACLs). The work also proposes Binder as an extension to the Datalog programming language, which is a subset of the Prolog language, in order to provide a formalised expressive security language. This shows the long history of logic usage for security-related problems and the variety of use cases which can be covered within the IT domain. In [54], an algebraic specification of the risk management process for IT systems is proposed. The authors use a scenario and attack tree-based security evaluation with the NetRAM approach, creating quantitative results for the support of associated stakeholders with regard to decision making. This work also considers automating the risk management process independently from the concrete implementation on an abstract level. A similar concept is defined by the author of [193] who proposes the MulVAL tool within Datalog to perform logic and rule-based attack simulation, policy checks, hypothetical analyses, and attack tree generation. The main contribution is an approach to perform security analyses based on the comparison of existing technical vulnerabilities and their violation of predefined policies, e.g. based on standards, guidelines, or company-internal documentation. In [204], an approach based on predicate calculus is described and a formal method is presented for the

analysis of the safety status of large-scale systems based on documented rules which are checked against the system status.

In general, rules are conditional sentences in a specific and formalised way which can be interpreted with logic, mimicking human reasoning in a suitable manner. A rule consists of a premise (if + condition) and a conclusion (then + deduction). When the premise is true, then the conclusion is also true. Figure 8.1 illustrates the general structure of rules (adapted from [44]). Rule-based expert systems shall only use conjunctions (AND relation), rules with disjunctions (OR relation) shall be split up into several conjunctive rules summarising the original statement [178]. Furthermore, rules can be integrated into expert systems in a forward-chained and a backward-chained way [178]. By chaining rules forwardly, conclusions are drawn based on given premises and the result is taken as a premise for the subsequent rule to move from starting conditions towards unknown conclusions. By chaining rules backwardly, possible causes for conditions are concluded. Following [44], OT security analyses, which include the security risk assessment processes, can best be modelled using forward-chained rules starting from the actual known state as a basis towards the desired goal, e.g. pending threats, associated security risks, or required countermeasures

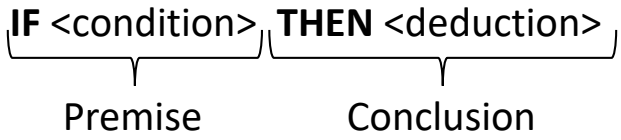

Figure 8.1: General structure of rules (adapted from [44])

8.2.2. Comparison of Logic Languages

The propositional logic is the most basic declarative language of logic available for the described problem. It contains only facts as an ontological commitment (representing the objects existing within the used language) for knowledge definition [58]. Nevertheless, due to the impossibility to define variables, objects, relations, and general statements, the propositional logic is not suitable for the problem at hand and was not further regarded. The current state of the art with regard to the necessary rules for the automated execution of security risk assessments show a clear tendency towards boolean decision making for the modelling of the described processes [46]. Therefore, the authors focused on the epistemological commitment (representing the possible types of information within the used language) of true, false, and unknown to define a semi-formal information model. The analysis of the security risk assessment process has shown that fuzzy logic is not required since there is no need for probabilities with a known interval value. All needed information have fixed values, mostly represented by enumerations or predefined scales without the requirement of covering imprecise

and non-numerical data points. Another alternative is the temporal logic, which is not required either since the described problem is lacking the necessity of a time aspect. This results in the choice of predicate logic as it is more expressive than propositional logic, providing, e.g. quantifiers and relations [44]. Predicate logic commits to the existence of objects and relations by additional expressive power, whereas propositional logic is only able to express facts [58]. The summary of logic characteristics can be found within Table 8.1.

Table 8.1: Comparison of available logic language (based on [58, 60])

Logic Language	**Ontological Commitment**	**Epistemological Commitment**
Propositional Logic	Facts	True / False / Unknown
Predicate Logic	Facts / Objects / Relations	True / False / Unknown
Temporal Logic	Facts / Objects / Relations / Time	True / False / Unknown
Fuzzy Logic	Facts with degree of truth	Known interval value

In summary, predicate logic (also called first-order logic) is the most suitable candidate out of the four available logic languages. It is favoured due to the usage of variables in contrast to fixed terms and the capability of drawing general and universally quantified conclusions. Another topic to be considered in this context is the possible utilisation of ontologies, e.g. with OWL which is also based on predicate logic, for the described information modelling requirements. Up to now, no reasoning capabilities to extract unknown knowledge within the information model are required, because everything that is needed is also modelled, representing an adapted closed world assumption. In addition, the amount of available representative instances to create an extensive ontology is too low when investigating the current use cases for the automation of security risk assessments and the specification of an ontology is therefore too resource-intensive in this case. Furthermore, the usage of predicate logic inherits a lower complexity, shorter development cycles, and less computational resources [58, 74]. Therefore, the current research direction does not include the usage of ontologies. Nevertheless, the usage of predicate logic ensures the possibility for a subsequent usage of ontologies because typically both models are interchangeable in their characteristics, such as contents and visualisation [58].

8.2.3. Logic-Based Rule Summary

This section summarises the logic-based rule definitions and the associated characteristics. Afterwards, the rules for each of the four security risk assessment phases of network segmentation (Section 8.2.4), requirements & guarantees (Section 8.2.5), risks (Section 8.2.6), and attestation (Section 8.2.7) are presented.

As shown before (Figure 8.1), each rule always consists of a premise representing the input(s) and a conclusion representing the output(s). Depending on the contents which are used within a premise or a conclusion, the specified rules from the subsequent sections can be categorised into Rule Classs (RCs). There are 15 rules in total which can be each aligned towards a RC. Table 8.2 shows the summary based on the

logical characteristics, such as required asset-related or security-related information. The text before the brackets describes the type of information and the text within the brackets refers to the specific content. Each line of the corresponding RC shows the examples from the distinctive security risk assessment phase.

Table 8.2: Summary of RCs based on the logical characteristics per security risk assessment phase (adapted from [44])

Rule Class	**Phase**	**Rule**	**Premise (Input)**	**Conclusion (Output)**
RC1	#1	1-1 & 1-2	Asset (Safety)	Security (Zones)
RC2	#1	1-3	Asset (Interface) + Security (Zones)	Security (Conduits)
	#3	3-1	Asset (Access Point) + Security (Vulnerability)	Security (Vulnerability)
	#3	3-2	Asset (Path) + Security (Vulnerability)	Security (Vulnerability)
	#3	3-3	Asset (Target) + Security (Vulnerability)	Security (Vulnerability)
RC3	#2	2-3 & 2-4	Security (SL)	Security (CR)
	#2	2-5 & 2-8	Security (SL)	Security (SR)
	#2	2-1 & 2-7	Security (SL)	Security (CR) + Security (SR)
	#2	2-2 & 2-6	Security (SL)	Security (SL)
	#3	3-4	Security (Vulnerability)	Security (Technique)

In general, all rules are written as complex sentences using the predicate logic and it is possible to define the needed decisions as definite logical formulas (Horn clauses) in the implication form (if-then-relationship) [187]. The following parts of the formal grammar of predicate logic are used [58]:

- Relations with varying arity (= amount of variables), e.g. $\boldsymbol{S(x)}$
- Variables (= terms), e.g. $\boldsymbol{x}$
- Different operators
  - And ($\wedge$)
  - Implication ($\Longrightarrow$)
  - Negation ($\neg$)
- Universal quantifier ($\forall \boldsymbol{x}$) or in a nested way ($\forall \boldsymbol{x} \forall \boldsymbol{y}$)

Each ZCR of the security risk assessment process (see Figure 5.6) contains a varying amount of rules which are performed in a sequential way. All these logic-based rules are created manually and do not need to be checked via, e.g. SHACL for conformity and correctness. This set of defined rules results in a huge decision tree based on the defined rules where every result influences the inputs for the subsequent rule in a chronological manner and where always only rule is active at a certain point in time. Therefore, the final security risk assessment result depends on the chain of rules and their individual decisions (= forward-chained [44, 178]). Nevertheless by doing so, the area of applicability and validity of the rules is focused on the information models from this dissertation and would need to be changed if were to be integrated into other approaches. Different application scenarios or use cases may require an adapted set of rules or a varying level of detail for the rules [44].

The presented usage of predicate logic enables the elicitation of security expert knowledge which is needed to perform security risk assessments [187]. This is based on the information model that was identified, collected, and formalised during the first three method steps. The combined result is a semi-formal information model of the described problem regarding security risk assessments. Moreover, the outcome is compliant to the definitions from the IEC 62443-3-2 security standard and addresses the requirements that were presented in Section 8.1.

8.2.4. Phase 1: Network Segmentation

The following logic-based rules specified as predicate logic are included within the first phase of network segmentation and are deducted from the swimlanes described in Section 6.2.2:

Rule 1-1: $\forall x \forall y \quad S(x) \land S(y) \land M(x,y) \implies A(x,y)$
Rule 1-2: $\forall x \forall y \quad \neg S(x) \land \neg S(y) \land M(x,y) \implies A(x,y)$
Rule 1-3: $\forall x \forall y \quad I(x,y) \land \neg Z(x,y) \implies C(x,y)$

Rule 1-1 shows the assignment of safety-relevant assets within the same module to a common security zone. Firstly, it is checked for both assets x and y if their SuitableForSafetyFunctions characteristic is true (S(x) ∧ S(y) == True). Secondly, it is checked if both assets are in the same module (M(x,y) == True), e.g. a laser engraving cell of a production process as within the presented demonstrator (see Section 5.1). Lastly, if S(x), S(y), and M(x,y) hold true, both assets are assigned into the same security zone (A(x,y) == True). Rule 1-2 refers to the same procedure but just for non-safety-relevant assets. Rules 1-1 and 1-2 belong to the RC1 from Table 8.2 containing asset-related information within the premise and security-relevant information within the conclusion.

Rule 1-3 represents the creation logic of the security conduits between the already defined security zones. At the beginning it is checked if the two assets x and y have a connected interface via a physical port (I(x,y) == True). Afterwards, both assets need to be in different zones to qualify for the assignment of a conduit following the IEC 62433-3-2 definitions (¬ Z(x,y) == True). The conclusion holds true if the premise is fulfilled. Then, both assets are assigned to a common conduit (C(x,y) == True), similar as with Rule 1-1 and Rule 1-2 mentioned before. Rule 1-3 can be categorised under RC2 from Table 8.2 combing facts from assets and security to conclude about additional security-relevant information.

The following fact sheets show the description of the used relations and their variables for the first phase of network segmentation for a better understanding:

```
Name: SuitableForSafetyFunctions
Description: Logical check if the given asset is
suitable for the usage within safety functions
based on the specified asset characteristic
(ECLASS IRDI 0173-1#02-BAD722#009).
```

```
Abbreviation: S
Arity: 1
Symbol: S(x)
Input (variables): Asset x
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset then
        if x.SuitableForSafetyFunctions == TRUE then
                return S(x) = TRUE
        else
                return S(x) = FALSE
else
        return MissingInformation
```

Listing 8.1: Fact Sheet SuitableForSafetyFunctions

```
Name: CheckForSameModule
Description: Logical check if the two given assets
are used within the same module based on the
specified asset characteristic (asset.Module).
Abbreviation: M
Arity: 2
Symbol: M(x,y)
Input (variables): Assets x and y
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset & y == asset then
        if x.Module == y.Module then
                return M(x,y) = TRUE
        else
                return M(x,y) = FALSE
else
        return MissingInformation
```

Listing 8.2: Fact Sheet CheckForSameModule

```
Name: AssignmentToSameZone
Description: Assignment of a security zone to the
two given assets and adaptation of the corresponding
asset characteristic (asset.Zone).
Abbreviation: A
Arity: 2
Symbol: A(x,y)
Input (variables): Assets x and y
Output (variables): TRUE/FALSE
Specification (as pseudo code):
```

```
if x == asset & y == asset then
        zoneName = createZoneName()
        x.Zone = zoneName
        y.Zone = zoneName
        return A(x,y) = TRUE
else
        return A(x,y) = FALSE
        return MissingInformation
```

Listing 8.3: Fact Sheet AssignmentToSameZone

```
Name: CheckForCommonInterface
Description: Logical check if the given assets
have a common (physical) interface based on the
specified asset characteristic (asset.PhysicalPort).
Abbreviation: I
Arity: 2
Symbol: I(x,y)
Input (variables): Assets x and y
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset & y == asset then
        if anyInterface(x.PhysicalPort, y.PhysicalPort)
                return I(x,y) = TRUE
        else
                return I(x,y) = FALSE
else
        return MissingInformation
```

Listing 8.4: Fact Sheet CheckForCommonInterface

```
Name: CheckForSameZone
Description: Logical check if the given assets
are within the same zone based on the specified
asset characteristic (asset.Zone).
Abbreviation: Z
Arity: 2
Symbol: Z(x,y)
Input (variables): Assets x and y
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset & y == asset then
        if x.Zone == y.Zone then
                return Z(x,y) = TRUE
        else
                return Z(x,y) = FALSE
```

```
else
        return MissingInformation
```

Listing 8.5: Fact Sheet CheckForSameZone

```
Name: AssignmentToSameConduit
Description: Assignment of a security conduit to the
two given assets and adaptation of the corresponding
asset characteristic (asset.Conduit).
Abbreviation: C
Arity: 2
Symbol: C(x,y)
Input (variables): Assets x and y
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset & y == asset then
        conduitName = createConduitName()
        x.Conduit = conduitName
        y.Conduit = conduitName
        return C(x,y) = TRUE
else
        return C(x,y) = FALSE
        return MissingInformation
```

Listing 8.6: Fact Sheet AssignmentToSameConduit

8.2.5. Phase 2: Requirements & Guarantees

The following logic-based rules specified as predicate logic are included within the second phase of requirements & guarantees and are deducted from the swimlanes described in Section 6.2.3:

Rule 2-1: $\forall c \forall s \quad SLCT(c) \implies CR(c, "Shifted") \wedge SR(s, "ToBeChecked")$
Rule 2-2: $\forall c \forall s \quad \neg SLCT(c) \implies SLAC(c)$
Rule 2-3: $\forall c \quad \neg SLAC(c) \implies CR(c, "Mitigated")$
Rule 2-4: $\forall c \quad SLAC(c) \implies CR(c, "ReconfigurationAdvised")$
Rule 2-5: $\forall s \quad SLCT(s) \implies SR(s, "Unmitigated")$
Rule 2-6: $\forall s \quad \neg SLCT(s) \implies SLAC(s)$
Rule 2-7: $\forall s \quad SLAC(s) \implies CR(c, "Mitigated") \wedge SR(s, "Mitigated")$
Rule 2-8: $\forall s \quad \neg SLAC(s) \implies SR(s, "Mitigated")$

All eight rules from this phase of the security risk assessment process can be categorised into the RC3 from Table 8.2 containing a premise based on security-related input and a conclusion resulting in additional security-related information. The premises always use the comparison of a combination of the three SL types

(capable, achieved & target) to determine the security status. Then, the conclusion deducts new knowledge about the SUC. Rule 2-4 functions here as a guiding and representing example. The CRs of each target asset are utilised for a check of the asset's SL-A in comparison to the SL-C. This is a check if the configured and achieved SL by the system integrator of the asset already has the maximum value limited by the capable SL defined by the component manufacturer during production. If this is the case (premise), it is concluded that the asset can be improved to increase the overall SL-A and therefore a reconfiguration is advised. Additional details about the underlying procedure can be found within the swimlane specification in Section 6.2.3.

The following fact sheets show the description of the used relations and their characteristics for the second phase of requirements and guarantees for a better understanding:

```
Name: ComparisonOfSecurityLevelsCapableAndTarget
Description: Logical check if the given component or
system requirement has a higher target security level
(SL-T) as compared to the capable security level (SL-C).
Abbreviation: SLCT
Arity: 1
Symbol: SLCT(r)
Input (variables): Component/System Requirement r
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if r == Requirement then
        if r.SL-C < r.SL-T then
                return SLCT(r) = TRUE
        else
                return SLCT(r) = FALSE
else
        return MissingInformation
```

Listing 8.7: Fact Sheet ComparisonOfSecurityLevelsCapableAndTarget

```
Name: ComponentRequirementStatusChange
Description: Change of the status of a given
component requirement, e.g. "Shifted", "Mitigated",
or "ReconfigurationAdvised".
Abbreviation: CR
Arity: 2
Symbol: CR(c, "X")
Input (variables): Component Requirement c \& Status "X"
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if c == CR then
        if c.Status ==  "X" then
```

```
                return CR(c, "X") = TRUE
        else
                return CR(c, "X") = FALSE
else
        return MissingInformation
```

Listing 8.8: Fact Sheet ComponentRequirementStatusChange

```
Name: SystemRequirementStatusChange
Description: Change of the status of a given
system requirement, e.g. "ToBeChecked", "Mitigated",
or "Unmitigated".
Abbreviation: SR
Arity: 2
Symbol: SR(s, "X")
Input (variables): System Requirement s \& Status "X"
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if s == SR then
        if s.Status ==  "X" then
                return SR(s, "X") = TRUE
        else
                return SR(s, "X") = FALSE
else
        return MissingInformation
```

Listing 8.9: Fact Sheet SystemRequirementStatusChange

```
Name: ComparisonOfSecurityLevelsAchievedAndCapable
Description: Logical check if the given component or
system requirement has a higher achieved security level
(SL-A) as compared to the capable security level (SL-C).
Abbreviation: SLAC
Arity: 1
Symbol: SLAC(r)
Input (variables): Component/System Requirement r
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if r == Requirement then
        if r.SL-A < r.SL-C then
                return SLAC(r) = TRUE
        else
                return SLAC(r) = FALSE
else
        return MissingInformation
```

Listing 8.10: Fact Sheet ComparisonOfSecurityLevelsAchievedAndCapable

8.2.6. Phase 3: Risks

The following logic-based rules specified as predicate logic are included within the third phase of risks and are deducted from the swimlanes described in Section 6.2.4:

Rule 3-1: $\forall x \quad A(x) \wedge CVSSAV(x, "Network") \implies CVSSS(x, "Changed")$
Rule 3-2: $\forall x \quad P(x) \wedge \neg CVSSAV(x, "Physical") \implies CVSSS(x, "Changed")$
Rule 3-3: $\forall x \quad T(x) \wedge \neg CVSSAV(x, "Physical") \implies CVE(x)$
Rule 3-4: $\forall x \quad CVSSS(x, "Changed") \implies CVE(x)$

Rule 3-1, 3-2, and 3-3 use a similar approach. All three rules are based on an initial check of the inspected asset and its associated type, namely access point, path, or target asset. Depending on the asset type, the CVSS attack vector characteristic is identified and used to determine if the present technical vulnerability can lead to a potential attack. For example, the access point assets represent the edges of a network (zone) and the interface of a network (conduit). Therefore, it is checked in the premise of Rule 3-1 if the CVSS attack vector is possible via the network level. Afterwards, Rule 3-4 is triggered and the CVSS scope characteristic is used to determine if the vulnerability can be used to impact other assets out of the original security scope resulting in a possibility for an attacker to proceed through the network and to further build up the kill chain. Then, the conclusion uses a predefined mapping[4] of MITRE ATT&CK techniques towards CVE values.

The following fact sheets show the description of the used relations and their characteristics for the third phase of risks for a better understanding:

```
Name: CheckForAcessPointAsset
Description: Logical check if the given asset represents an
access point based on the asset characteristic
   (asset.AccessPoint).
Abbreviation: A
Arity: 1
Symbol: A(x)
Input (variables): Asset x
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset then
        if x.AccessPoint == TRUE then
                return A(x) = TRUE
        else
                return A(x) = FALSE
else
        return MissingInformation
```

Listing 8.11: Fact Sheet CheckForAcessPointAsset

[4] https://github.com/center-for-threat-informed-defense/attack_to_cve

```
Name: CheckCVSS-AttackVector
Description: Logical check if the given vulnerability
represented by a CVE entry matches the given attack
vector value, e.g. "Network", "Adjacent", or "Local".
Abbreviation: CVSSAV
Arity: 2
Symbol: CVSSAV(x, "X")
Input (variables): Asset x \& Value "X"
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset then
        if x.Vulnerability.AttackVector == "X" then
                return CVSS-AV(x, "X") = TRUE
        else
                return CVSS-AV(x, "X") = FALSE
else
        return MissingInformation
```

Listing 8.12: Fact Sheet CheckCVSS-AttackVector

```
Name: CheckCVSS-Scope
Description: Logical check if the given vulnerability
represented by a CVE entry matches the given scope value,
e.g. "Changed".
Abbreviation: CVSSS
Arity: 2
Symbol: CVSSS(x, "X")
Input (variables): Asset x \& Value "X"
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset then
        if x.Vulnerability.Scope == "X" then
                return CVSSS(x, "X") = TRUE
        else
                return CVSSS(x, "X") = FALSE
else
        return MissingInformation
```

Listing 8.13: Fact Sheet CheckCVSS-Scope

```
Name: MapTechniquesToCVE
Description: Matching of the given techniques
with the present vulnerabilities (to develop a
kill chain as a possible attack path).
Abbreviation: CVE
Arity: 1
```

```
Symbol: CVE(x)
Input (variables): Asset x
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset then
        if match(x.vulnerabilities, x.techniques) then
                return CVE(x) = TRUE
        else
                return CVE(x) = FALSE
else
        return MissingInformation
```

Listing 8.14: Fact Sheet MapTechniquesToCVE

```
Name: CheckForPathAsset
Description: Logical check if the given asset represents an
path asset on the asset characteristic (asset.Path).
Abbreviation: P
Arity: 1
Symbol: P(x)
Input (variables): Asset x
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset then
        if x.Path == TRUE then
                return P(x) = TRUE
        else
                return P(x) = FALSE
else
        return MissingInformation
```

Listing 8.15: Fact Sheet CheckForPathAsset

```
Name: CheckForTargetAsset
Description: Logical check if the given asset represents an
target asset on the asset characteristic (asset.Target).
Abbreviation: T
Arity: 1
Symbol: T(x)
Input (variables): Asset x
Output (variables): TRUE/FALSE
Specification (as pseudo code):
if x == asset then
        if x.Target == TRUE then
                return T(x) = TRUE
        else
```

```
                return T(x) = FALSE
else
        return MissingInformation
```

Listing 8.16: Fact Sheet CheckForTargetAsset

8.2.7. Phase 4: Attestation

The fourth phase of attestation does not contain any logic-based rules specified within predicate logic as fact sheets for explanation, because it only describes the collection and documentation of information from the previous three phases for the sake of documentation. There are no decisions based on logic necessary because all tasks within this phase are fully automated without the need for specific security expert knowledge.

8.3. Intermediate Summary

The third method step of information usage further enhances the collected and formalised information base towards the automation of security risk assessment processes. The semi-formal information model from within the swimlanes is taken here as a basis for the definition of logic-based rules to elicit the inherited expert knowledge and the decision-making processes. This addresses the third deficit D3 of low approach maturity from Section 4.3.3. First, the associated requirements for this method step are specified and used to introduce and motivate the usage of rules. Afterwards, a comparison of the available logic languages, especially propositional, fuzzy, temporal, and predicate logic, is performed. Due to the ontological and epistemological commitment of the predicate logic, this type of logic is chosen for the subsequent definition of rules. The security risk assessment phases of network segmentation, requirements & guarantees, and risks are modelled in distinctive predicate logic rules according to the respective swimlane contents. The attestation phase does not contain any logic and therefore does not have any rules defined. Figure 8.2 shows the overall dissertation structure and the current progress regarding the third method step of information usage.

Figure 8.2: Dissertation progress after the third method step of information usage

This method step of information usage aligns with the general knowledge engineering process step of encoding expert knowledge from [58]. To perform the remaining method step towards the automation of security risk assessments, the information access needs to be defined to be able to support the specified information model with the necessary real-world information representing the SUC. Therefore, the next and final method step within the upcoming chapter motivates the usage of DTs, especially AASs for the industrial domain, and presents the necessary concepts to fulfil the demands of the proposed solution approach.

9. Information Access

The results of the previous three method steps of information collection, formalisation, and usage represent the theoretical basis for the automation of security risk assessments. This is achieved via the identification and analysis of required information and the specification of security expert knowledge within a model and the associated decision logic rules. To complete the proposed solution approach, a practical definition of an interface including a communication channel to access the information necessary for the formalised security risk assessment is presented. This will enhance the overall credibility and usability. Furthermore, it opens up the possibility to prototypically implement and practically evaluate the resulting LOA. The overall aim is to integrate an up-to-date and future-proof concept for data access into the overall solution approach. The remaining requirements are identified and analysed within Section 9.1. The rest of the chapter includes the introduction, motivation, and comparison of the DT concept and the available technological approaches to implement it. The AAS, OPC UA, AML, and the Module Type Package (MTP) are further investigated due to their relevance within Section 9.2. Afterwards in Section 9.3, the specific translation of the developed UML information model into the AAS structure via three newly self-developed AAS submodel templates is presented. In the end, Section 9.4 summarises the contents of this chapter and reflects the overall progress.

9.1. Requirements Analysis: Information Access

This section contains the fact sheets for requirements resulting from the related work analysis and the literature research regarding the automation of security risk assessments, representing the method step of information access (abbreviated with "IA"). Each requirement (abbreviated with "R") can be identified via its short name as the title in bold and a unique identifier in the following style "R-IA-#". Furthermore, a description of the requirement is given to summarise the demanded contents. The three different verbs which are used, represent the relevance of each requirement in a prioritising order (in accordance to the RFC 2119[1] and the ISO/IEC Directives[2]): High = *"shall"* / Medium = *"should"* / Low = *"may"*. All requirements are not directly measurable and are therefore evaluated in a qualitative manner based on a subjective justification by the author of this dissertation later on. Finally, associated

[1] https://datatracker.ietf.org/doc/html/rfc2119
[2] https://www.iso.org/sites/directives/current/part2/index.xhtml

and generally covered statements from other references are listed to further support the requirements analysis.

Requirement: OT Adaptation

- ID: R-IA-1
- Description: The access to the relevant information regarding the security risk assessments *shall* be adequate for the OT domain.
- Generally covered statements from other references:
  - The specifics of IACSs shall be regarded, e.g. availability, real-time characteristics, safety, or regulative demands [44, 74].
  - The information retrieval shall be minimally invasive being based on passive approaches [74].
  - The main application is the mirroring of actual industrial components during their usage [205].
  - A possible solution shall regard the low amount of available resources for computing on industrial components [74].
  - Support of the various lifecycle phases of industrial components [206].
  - Integration of standardised communication interfaces enabling a uniform information access to various types of data relevant for the security risk assessments [74, 206, 207].
  - Possibility to integrate, extend, and map already existing information models to achieve a high domain coverage [208].
  - It is essential to integrate security-relevant information into machine-readable, machine-processable, and machine-comprehensible models [32].
  - A digital data model is one of the main pillars for security-related decision making [184].

Requirement: Syntax & Semantics

- ID: R-IA-2
- Description: The information access *shall* use standardised syntax and semantics to achieve a high credibility and usability.
- Generally covered statements from other references:
  - A prerequisite is the digital representation of the physical world within the information world available whenever and wherever it is needed [209].
  - Integration of component information in various degrees of detail and type [44, 206, 210].

  - Common and generic definitions are needed to achieve a high interoperability and interchangeability of the formalised information [209].
  - Concepts to describe complex relations and constraints within the modelling concept to achieve a high expressiveness [208].
  - The specified information shall be traceable and auditable [210].
  - Possibility to specify stand-alone information models for certain domains, use cases, or application areas [206].
  - Storage of information models within predefined libraries for better distribution, access, and reuse [183].
  - The solution should have a high persistence resulting in suitability for storing, retrieving, and querying models [208].

Requirement: Secure Access

- ID: R-IA-3
- Description: The access to the information *should* be secured via state of the art technological implementations.
- Generally covered statements from other references:
  - Access rights and rules to the information should be reduced to authorised entities to protect confidential knowledge and prevent manipulation [44].
  - The transmission of security-relevant information should ensure confidentiality, integrity, and authenticity [74].
  - Sources of knowledge should be checked and provide versioning to increase user trust [44].
  - A mature level of security should be regarded for real-world implementations and applications [211].
  - Secure access to information should reduce the possible impacts due to a single point of failure [207].
  - Security feature support, e.g. authentication, authorisation, or usage control [212].

Requirement: Tool Availability

- ID: R-IA-4
- Description: The concept for information access *should* have implemented and tested software tools available.
- Generally covered statements from other references:

- The quantity and quality of elaborated tools with suitable Application Programming Interfaces (APIs) should be given to enable the usage of up-to-date approaches, concepts, and technologies [208].
- Visualisation should support the tool usage for the different users including an adequate GUI [211].
- The possible solution should have a high community acceptance including stakeholder activity, regular updates, and the usage for various domain use cases [208].
- Open-source, free of charge, and license-free software tools should be used for the digitalisation of manufacturing and I4.0 [211].
- Available software tools are one of the main pillars for security-related decision making [184].

Requirement: Data Sources

- ID: R-IA-5
- Description: The information access *may* be based on various types of data sources to feed the security risk assessment process.
- Generally covered statements from other references:
  - Support of heterogeneous data sources for integrating knowledge with varying maturity and access mechanisms [74].
  - The integration of data may ensure the highest possible degree of automation via standardised communication interfaces [44].
  - Adaptation to changing and flexible data sources [74].
  - The imported data may be as up-to-date as possible to guarantee a correct information based for the security risk assessment process [44].

9.2. Introduction to Digital Twins

The term DT and the associated definition were first coined by the researcher Michael Grieves from the University of Michigan in 2003. The National Aeronautics and Space Administration (NASA) picked them up in 2010 within one of the drafts for their technology roadmap [205, 212]. The original idea was to have an integrated simulation of a real-world system to mirror the characteristics and the behaviour in the digital world, enabling the application of simulation-based systems engineering [205]. At that time, the fulfilment of the DT was foreseen to happen roughly in 2027 [205], but looking at the topic and the current developments now, we are in the middle of the first stable DT implementations and practical use [213]. Therefore, there are various conceptual and technological approaches currently available that aim to fulfil the DT concept [213].

Furthermore, the usage of DTs is foreseen to transform the currently fragmented security modelling landscape [213]. Nowadays, there are isolated information silos present operated by third parties and based on commercial, individual, proprietary, or custom-built security models, such as taxonomies or ontologies [213]. Therefore, this dissertation utilises the DT concept as a unified, virtual, and standardised representation of relevant information regarding security risk assessments for the OT domain [213]. By doing so, every industrial stakeholder, e.g. the component manufacturer or the system integrator, is enabled to understand and use DTs for information modelling without any additional third party being involved. In addition, the scope of this dissertation is set towards security engineering using DTs [79], not towards the development of mechanisms to secure the DT as for example in [207, 210, 214] as this is always tied to a certain implementation and its characteristics [211]. Consequently, the main research results are within the topic of information access for the automation of security risk assessments using DTs, not within the further improvement of the DT and its security capabilities itself. Therefore, the following subsections will only investigate the available conceptual and technological DT approaches which can be utilised for the described purpose of this dissertation.

The research and development landscape regarding DTs is very diverse and in general, a plethora of approaches is available. Due to their attention, publicity, distribution, and development extent in Europe, the AAS, OPC UA, and AML are investigated further and compared [206]. In addition, a concurrent analysis of the MTP from the process industry domain is performed [215]. Various minor approaches, such as The Network Configuration Protocol (NETCONF) together with Yet Another Next Generation (YANG) [208] or OWL [208], which can be adapted towards a DT use, are neglected here in the further analysis.

9.2.1. Asset Administration Shell

The first approach towards an industrial DT implementation is the AAS. In 2016, the Plattform Industrie 4.0 (PI4.0)[3] from Germany started the I4.0, RAMI4.0, and AAS activities as a technology network consisting of actors from companies, organisations, science, and politics. Since 2021 the Industrial Digital Twin Association (IDTA)[4] based in Germany took over the AAS developments as a responsible entity [211]. The AAS is a further building block within the overall interoperability landscape by providing a digital representation of industrial assets, such as components, machines, plants, software, or even documents [206]. The combination of an AAS and the associated physical asset is called an I4.0 component [207]. As huge advantages, it inherits standardised information models based on submodels[5] for, e.g. nameplates, contact information, or hierarchical structures enabling a BOM, uniform data access mechanisms in an organising and aggregating role throughout the complete asset lifecycle [212, 216, 217]. Furthermore, the AAS enables a minimally invasive and

[3]https://www.plattform-i40.de/IP/Navigation/EN/Home/home.html
[4]https://industrialdigitaltwin.org/en
[5]https://industrialdigitaltwin.org/en/content-hub/submodels

resource-saving way to access the asset information without affecting the physical system [213]. Since 2022, the AAS specification is standardised within the IEC 63278. The current security status of the AAS is twofold: Firstly, technical implementations are available regarding controls, such as access control, user rights and roles definitions, and integrity checks, from the research domain [207, 210, 214] or from the AAS specification itself[6]. Secondly, applications using the AAS to increase the security or the safety of a given system are still under development, but are considered, e.g. by IDS or IPS [88, 218, 219] or security and safety risk assessments [41, 207]. Nevertheless, the modelling and integration of AASs should be done carefully due to a possible single point of failure and the addition of new attack vectors to the underlying system [213]. In general, the alignment and mapping of the AAS towards other approaches, such as AML or OPC UA, are currently ongoing [206, 208]. The results from these discussions will define and impact the future developments of the AAS. More details about the AAS can be found within the subsequent Section 9.3.

9.2.2. OPC UA

The second approach that is relevant here is OPC UA which represents a holistic collection of technologies for the domain-independent semantic description of industrial information and the secure exchange of data during operation [206]. In addition, OPC UA offers platform independence, non-proprietary communication, and a high scalability. It was first published in 2006 as a service-oriented architecture and has since been further developed by the Open Platform Communications (OPC) Foundation[7], currently resulting in the IEC 62541 standard [206]. OPC UA offers possibilities to transform the classical automation pyramid into a modern information network, following the general principles of the I4.0. The basis is the OPC UA metamodel which provides defined syntax and semantics as a key factor for a vendor-independent digital data exchange [206]. Furthermore, the associated companion specifications exactly specify what kind of information are exchanged to fulfil a certain demand or use case, e.g. for securing the communication[8] or a special industrial fieldbus[9]. The current implementations of OPC UA use various known communication technologies for IT and OT applications in either client/server or publish/subscribe models, e.g. Transmission Control Protocol (TCP), User Datagram Protocol (UDP), Message Queuing Telemetry Transport (MQTT), or Hypertext Transfer Protocol (HTTP) Representational State Transfer (REST) interfaces [206, 208]. In addition, OPC UA follows the security by design approach, offering different levels of security for the accessibility and transmission of information [206].

[6] https://admin-shell-io.com/screencasts
[7] https://opcfoundation.org
[8] https://reference.opcfoundation.org/Core/Part2/v105/docs
[9] https://opcfoundation.org/developer-tools/documents/?type=specification

9.2.3. AutomationML

AML is the third approach investigated to fulfil the requirements of an industrial DT implementation. It has been developed since 2006 by the German AutomationML e.V.[10] as a flexible data modelling language for the storage and exchange of object models as files within engineering tool chains, e.g. assets, hardware and software topologies, or networked system models [206]. The AML e.V. supports users with implementations, knowledge, and training regarding the contents standardised within the IEC 62714 [206]. In general, AML is based on XML and offers an open, neutral, and free way to exchange data between industrial applications and tools, focusing on the engineering phase of a system lifecycle [220]. For AML to be suitable towards a wide range of applications, it is enhanced with domain libraries, e.g. for factory automation or the oil/gas process industry[11], supplementing the standard syntax and semantic with the definition of information from dedicated use cases [206]. The security-related utilisation of AML is currently evolving as already shown in Section 4.2 which refers to AMLsec [32, 55, 146] as a special modelling methodology for IACSs. In AMLsec, there are various libraries available which include relevant information for security risk assessments and the associated SUC [32]. Another example of using AML to specify hazard and vulnerability libraries for risk assessments (in this case a coupling of safety and security) is given in [221].

9.2.4. Module Type Package

The last approach that is relevant in the context of industrial DTs comes from the modularisation of process plants. Originally, the MTP concepts were documented within the NAMUR NE 148 "Automation Requirements relating to Modularisation of Process Plants" in 2013. Afterwards, the NAMUR[12] (an international association of user companies from the process industry that represents their interests concerning automation technology) proceeded with the development of the MTP and other organisations joined the efforts, resulting in the VDI/VDE/NAMUR 2658 from 2019 and the VDI 2776 from 2020. Just recently, the responsibility regarding the MTP was transferred to the PROFIBUS & PROFINET International (PI) organisation[13]. The combination of process and automation engineering regarding the MTP is currently best described within the recommendation of actions in [215]. The MTP itself is a file-based information model using AML to describe general characteristics of system module types, e.g. services, interfaces, or communication capabilities, which are shared between specific instances [216]. In addition, the Process Orchestration Layer (POL) is used during the engineering phase to configure specific module instances, which are then called Process Equipment Assemblys (PEAs) and represent procedural steps, such as distillation, consisting of at least one Functional Equipment

[10] https://www.automationml.org
[11] https://www.automationml.org/industrial-application/domain-model
[12] https://www.namur.net/en/about-us.html
[13] https://www.profibus.com/technologies/mtp

Assembly (FEA) which forms a process engineering function, such as heating or piping. By doing so, a specific module topology reflecting the physical layout of the process can be engineered and complex process plants can be created in a faster time to market [6]. To generally perform security analyses with the MTP has not yet been considered in an adequate manner and this is why the German VDI/VDE-GMA FA 3.22 working group was established in 2022[14]. In addition, within the PI, there are ongoing security-related activities regarding the MTP as well. In contrast, the topic of safety is already regarded within the MTP concepts as a basic requirement [6]. Currently, the research focus is shifting towards integration of the MTP into the AAS to further align the various available approaches [216]. Nevertheless, the MTP is not considered any further from here on due to the lack of security-related modelling capabilities and the narrow scope within the process industry.

9.2.5. Comparison of Digital Twin Approaches

The remaining three approaches of AAS, AML, and OPC UA are already integrated into practical applications from various domains [206]. Nevertheless, they are still subject to change and further developments [206]. In addition, discussions are ongoing to clarify the alignment of all three approaches in order to create an appropriate interoperability for the industrial domain [208]. The current status envisions the AAS as a centrally accessible data hub for information regarding the SUC in a predominantly organising and aggregating role [206]. In addition, everything which is engineering-related should be covered by AML and aspects regarding communication during system runtime should be solved by OPC UA [206]. To further improve the various building blocks with regard to interoperability, certain recommended actions were published, e.g. to avoid uncontrolled growth of AAS submodels, AML domain models, and OPC UA companion specifications, to reference between approaches, to reduce re-modelling and duplications, to design an easy data access, and to implement a targeted and economical data management [206]. Based on the provided comparison within the subsections before, this dissertation focuses on the AAS and its associated concepts from here on due to the following summarised reasons (based on Section 9.2.1 and Section 10.1):

1. OT Adaptation (R-IA-1): The AAS is developed specifically for the exchange of engineering data between the OT and IT domains and the associated requirements, e.g. minimally invasive data access, complete lifecycle support, and resource-saving implementations. In addition, the syntax and semantics are well-defined and provide a standardised information model, data access, and communication capabilities. The architectural design of using submodels enables the integration of and alignment with other information models as well.

[14]The author of this dissertation is actively taking part within the VDI/VDE-GMA FA 3.22 working group.

2. Syntax & Semantics (R-IA-2): Due to the current level of developments, the AAS enables the formalisation of various degrees of information, e.g. generic (type) data and specific (instance) data. This results in a high expressiveness and persistence. All information are enhanced with interoperable semantic IDs which enable a traceable and auditable dictionaries. The possibility to define specific submodels for certain use cases, domains, or applications and the availability of predefined formalisations make the AAS a suitable candidate for the overall information access.

3. Secure Access (R-IA-3): Available implementations of the AAS provide a mature level of security, e.g. access control, user rights and roles management, and certificate-based authentication and authorisation, and ensure confidentiality, integrity, and authenticity. Nevertheless, the achieved level of security is always dependant on the correct configuration and implementation of the provided software and the associated users. In an unfavourable scenario, the AAS (but also OPC UA, AML, or MTP) could also turn into a single point of failure resulting in a security risk for the stored information and knowledge.

4. Tool Availability (R-IA-4): The tool landscape for the AAS is broad and mainly based on research projects and the associated consortia. More details can be found within Section 10.2. Nevertheless, the majority of available approaches is open-source, free of charge, and license-free, resulting in a high acceptance among the community. Certain implementations, e.g. the AASX Package Explorer (see Section 10.1), also provide a useful GUI to increase user acceptance and usability.

5. Data Sources (R-IA-5): The AAS defines standardised data formats, interfaces, and communication functionalities. This enables the integration of different and heterogeneous data sources in an automated way. In addition, the direct connection to the underlying assets ensures an up-to-date information base and permits a regular verification of the data quality, e.g. by versioning.

The next section will focus on the utilisation details of the AAS and what conceptual aspects are used to fulfil the fourth method step of information access.

9.3. Utilisation of the Asset Administration Shell

The aim of this section is to describe the transformation process of the previously created UML class diagrams from Section 7.2 for the method step of information formalisation. The information from the UML class diagrams is transformed into the defined data syntax and semantics of the AAS for this method step of information access. Therefore, two information modelling aspects of the AAS need to be regarded [74]: (1) The AAS metamodel specification provided by the IDTA [222] and (2) the domain-, usage-, or technology-specific submodels [222]. The AAS

metamodel specification includes the technology-neutral information model defined in UML and it builds the foundation for the translation into various other formats for data exchange, such as XML, JavaScript Object Notation (JSON), RDF, AML, or OPC UA [222]. In addition, the submodels can be used to create a tree-like structure of information stored within an AAS. Each submodel can again be divided into submodel elements, e.g. properties containing single values of various data types, references to other AAS elements, submodel element collections, or even files [222]. Every element of this structure, such as an AAS itself or the subordinate submodels, has a unique ID for identification which can be assigned in a custom way [222]. Furthermore, it is recommended to reference external sources via semantic IDs as much as possible to increase expressiveness and interoperability, e.g. via Common Data Dictionary (CDD), Internationalized Resource Identifier (IRI), or IRDI [220]. By doing so, the AAS offers a standardised data structure and makes the stored information searchable and explorable [206]. The specification of the AAS-specific UML modelling which follows the IDTA metamodel is presented in this section. Afterwards, the specific AAS submodel template definition and development is presented within Section 10.2.

The IDTA offers 85 registered AAS submodel templates (as of 06. February 2024) covering various domains and use cases[15], ranging from technology-specific submodels, such as OPC UA or MTP, to more abstract submodels, such as maintenance, programming, or monitoring, to component-specific submodels, such as nameplates, technical data, or documentation, and to topic-related submodels, such as AI, safety, or security engineering. The corresponding status of the IDTA-registered submodels looks as follows, showing a clear tendency towards submodels being work in progress, but might be subject of changes and updates:

- Published: 20
- In review: 1
- In development: 41
- Proposal submitted: 21
- On hold: 2

Several already registered and published submodel templates are referenced and integrated into the results of this dissertation, e.g. the nameplate for asset information. The complete overview is provided later on within this section. Nevertheless, a fitting published submodel template for the information regarding the automation of security risk assessments is currently missing completely. Only a similar approach towards a safety submodel is also conducted within the BaSySafe[16] research project[17]. However,

[15]https://industrialdigitaltwin.org/en/content-hub/submodels
[16]https://www.dfki.de/web/forschung/projekte-publikationen/projekt/basysafe
[17]https://www.softwaresysteme.dlr-pt.de/media/content/Projektblatt_BaSySafe_01IS21015.pdf

there is one active working group specifying a submodel template for the modelling of information regarding the security engineering of process automation plants. This submodel template is to be developed by the NAMUR AK 1.3 security working group[18], but there are no results available yet to be reused [57]. A more detailed view on the general security engineering information model from the respective working group can be found in Section 4.3.

Therefore, this dissertation proposes a submodel template for the defined scope towards the automation of security risk assessments. The results are not yet registered at the IDTA as a submodel template for AASs[19] and have to be regarded as a status open for discussion, feedback, and further improvements. A submodel template represents a high-level blueprint of a submodel, which is used later on to create specific submodels for asset instances [222]. By doing so, it is possible to store the required information about each asset needed for the automated security risk assessment within the AAS in a machine-readable format [57]. Afterwards, the achieved results can be aligned with the submodel template in development towards the topic of security engineering to achieve a high level of reusability.

The basis for the AAS submodel template specification are the already defined UML class diagrams from Section 7.2, originally describing the needed information and the associated data types. Figure 9.1 shows the overall summary of the AAS regarding the SUC and the corresponding submodels which are in use. The naming of the submodels is based on their respective idShort which can be freely assigned during the submodel template definition.

The three white submodels (described by "Digital Nameplate for Industrial Equipment", "Contact Information", and "Hierarchical Structures enabling BOM") are already published[20]. The integration into the approach of this dissertation is further described towards the end of this section. In addition, the three black submodels are newly developed submodel templates originating from this dissertation and the associated research activities. Their characteristics are explained as follows:

- Main submodel: SecurityRiskAssessment
- Supporting submodel: SecurityLevelIEC62443
- Placeholder submodel: Miscellaneous

The overall architecture consists of three hierarchical levels of AASs forming the SUC: (1) Machine, (2) Module, and (3) Component. On the top level of the AAS hierarchy, there is an AAS of the machine referring via the hierarchical structures submodel to modules described by their AASs [57]. The AAS of the machine also contains the nameplate submodel for identification and the contact

[18]The author of this dissertation is actively taking part within the NAMUR AK 1.3 security working group.

[19]https://industrialdigitaltwin.org/en/wp-content/uploads/sites/2/2022/01/2021-12-01_IDTA_Process-Submodels_V1.0.pdf

[20]https://industrialdigitaltwin.org/en/content-hub/submodels

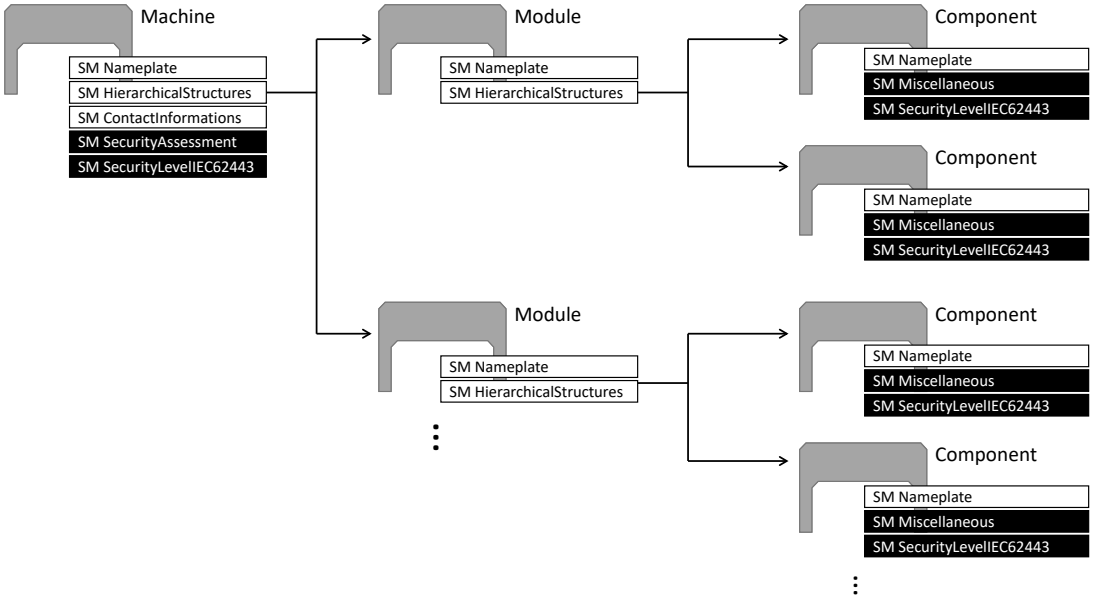

Figure 9.1: Summary of the AAS structure and relations

information submodel to define, e.g. the responsible stakeholders. In addition, the SecurityRiskAssessment submodel and the SecurityLevelIEC62443 submodel are used within the AAS of the machine and contain security-relevant data which are referenced for the complete machine as outputs of the security risk assessment process, e.g. zones and conduits, the SLs, or the security risk assessment results in the form of the attestation document. The AAS of each module uses the nameplate and hierarchical structures submodels, in this case not containing any security-relevant data. At last, the components are described by AASs including the already published nameplate submodel and the two self-developed submodels Miscellaneous and SecurityLevelIEC62443 which include the inputs for the security risk assessment process, e.g. capable and achieved SLs, present technical vulnerabilities, or the physical network topology. To avoid the duplication of information and to limit the nesting of submodel elements, reference elements are used between the different AASs from the three hierarchical levels [57]. They can refer to any element ranging from a complete AAS to a single property via the unique ID, e.g. as IRI, IRDI, or idShort. By doing so, duplicate information can be avoided by inheriting them enabling chaining and jumping and less modelling has to be done, e.g. in contrast to the usage of relationships. This results in a split of the information needed for or created by the automated security risk assessment process onto the three levels of the AAS with regard to the machine, the module, or the component. The distribution and distinction of AASs and their corresponding submodels is the result of long

discussion and should represent the current style of best practices in alignment towards the IDTA guideline on how to create a submodel template specification[21].

After the provision of an overview of the hierarchical AAS structure and the associated usage of submodels, the three newly developed submodel templates are further described here in a prioritised order. Each corresponding figure (Figure 9.2, Figure 9.3, and Figure 9.4) was semi-automatically created by using the combined functions of the AASX Package Explorer[22] and the open source PlantUML tool[23].

The first and most important submodel template is the SecurityRiskAssessment. It is only found on the machine level of the hierarchical AAS structure and fulfills a summarising function for the security risk assessment results. Figure 9.2 provides an overview of the included submodel element collections, references, and properties. The three main elements of zones, conduits, and attestation form the basis of the information. The other submodel elements refine and enhance these elements with further details, e.g. by asset information, accountable and responsible stakeholders, and results from the attestation.

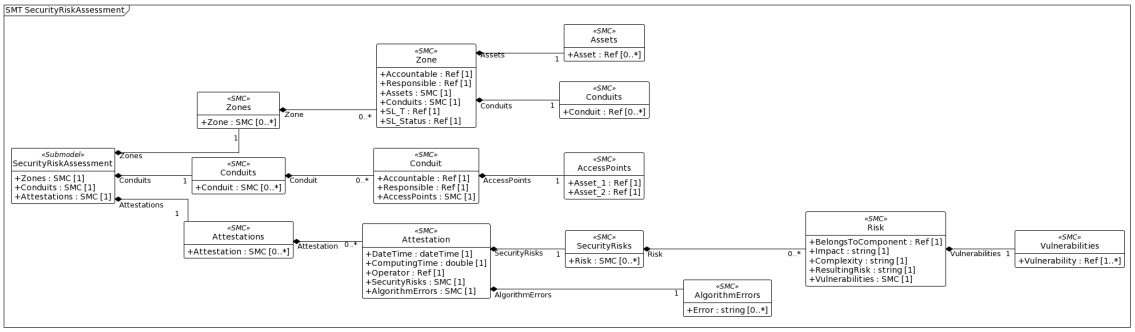

Figure 9.2: UML class diagram of the SecurityRiskAssessment submodel

The second submodel template is used in a supporting manner and is further shown in Figure 9.3. The SecurityLevelIEC62443 submodel contains one submodel element collection for the blueprint of an SL with seven submodel element collections to represent the seven FRs with their varying amount of properties for the CRs and SRs from the IEC 62443 standard. By doing so, a translation from the text-based description of the standard towards the machine-readable documentation within the AAS submodel is achieved and the information can be used for the automated security risk assessment process. Later on during the practical usage, various submodels are instantiated in order to cover the three different types (capable / achieved / target) of SLs. This is enabled by the AAS characteristic of multiplicity, which is defined here as zero to many.

Figure 9.4 shows the Miscellaneous submodel template which includes three different kinds of information relevant for the security risk assessment process which

[21] https://industrialtwin.org/en/wp-content/uploads/sites/2/2022/12/I40-IDTA-WS-Process-How-to-write-a-SMT-FINAL-.pdf

[22] https://github.com/eclipse-aaspe/aaspe

[23] https://plantuml.com/en

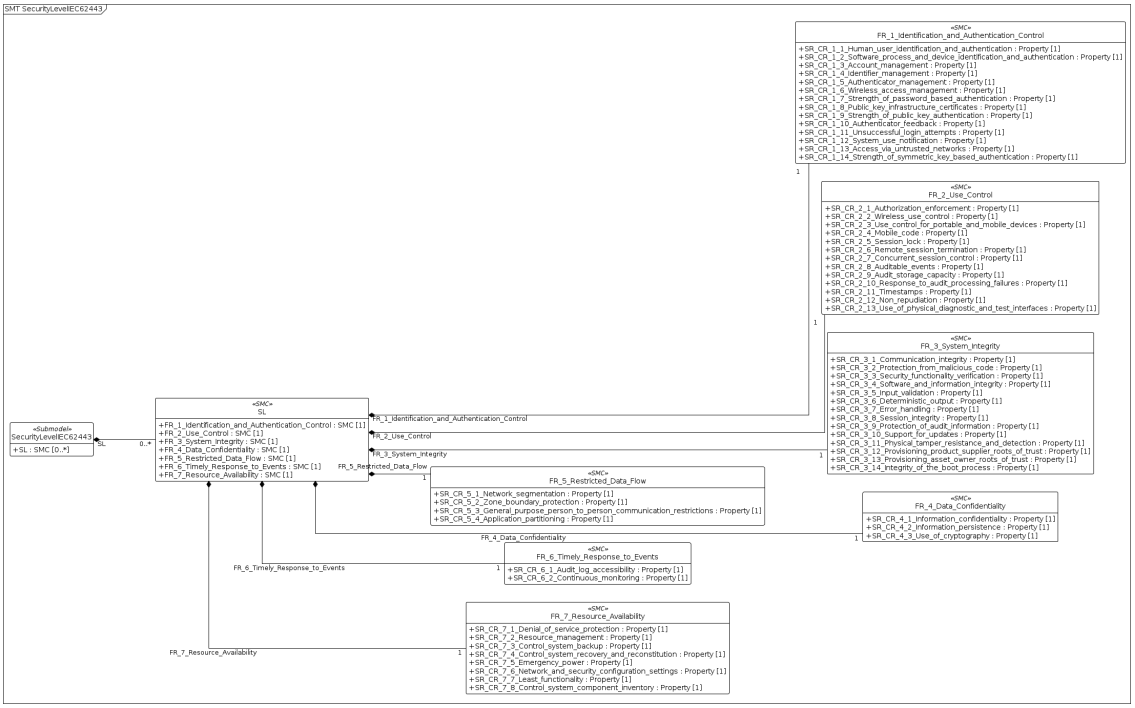

Figure 9.3: UML class diagram of the SecurityLevelIEC62443 submodel

are stored in a distinctive submodel element, on the component level of the hierarchical AAS structure: (1) the safety characteristic of an asset based on the ECLASS property “SuitableForSafetyFunctions” with the IRDI "0173-1#02-BAD722#009" as a boolean property, (2) the inherited technical vulnerabilities of the asset stored as CVE values within a submodel element collection derived from, e.g. external databases or finding reports, and (3) the physical network topology represented by the used network ports of the asset that are specified within a submodel element collection. This submodel is a temporary placeholder for necessary information which are currently not yet integrated in any other AAS submodel. Following the future developments of AAS submodels, the information from the Miscellaneous submodel could be superseded by the following submodels:

1. Safety characteristics → Upcoming submodel from the AutoS^2 research project[24] or the IDTA submodels[25] "Functional Safety" and "Reliability"

2. Technical vulnerabilities → "Vulnerability Management" submodel[26] from Interopera currently in development, in which an alignment via the AutoS^2 research project[24] took place, e.g. by implementing Vulnerability Exploitability eXchange (VEX)

[24]https://www.init-owl.de/forschung/projekte/detail/automatische-bewertung-und-ueberwachung-von-safety-security-eigenschaften-fuer-intelligente-technische-systeme
[25]https://industrialdigitaltwin.org/en/content-hub/submodels
[26]https://interopera.de/wp-content/uploads/2022/10/2022_07_Vulnerability_Management.pdf

3. Network topology → "Hierarchical Structures enabling BOM" or "Security Engineering" submodels from the IDTA[25] or the developments around the Software Bill of Material (SBOM) in general

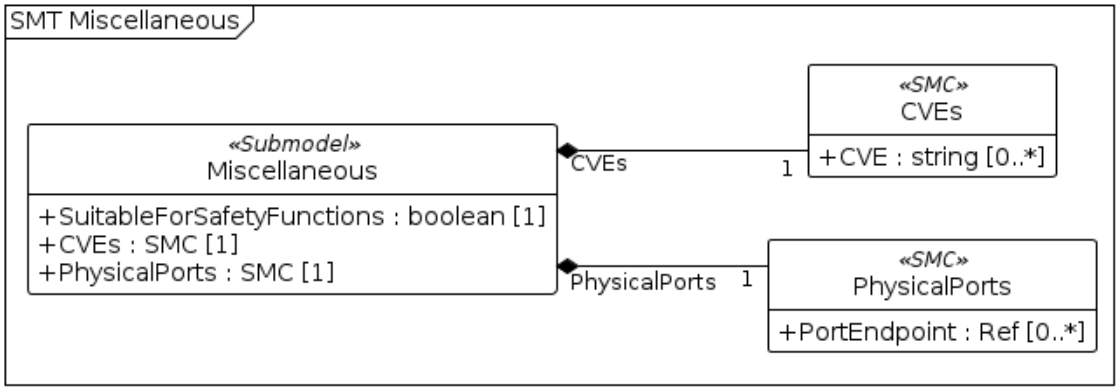

Figure 9.4: UML class diagram of the Miscellaneous submodel

In addition to the three self-developed submodel templates, three other submodel templates were used as shown in Figure 9.1. The published "Digital Nameplate for Industrial Equipment" submodel[27] is used on every AAS hierarchical level to store specific asset information, such as manufacturer name or the serial number of the component. By using the published "Contact Information" submodel[28] within the machine AAS, it is possible to define the roles of the operator and the responsible and accountable stakeholders in a single point avoiding duplications. The original intent of the IDTA is to address service issues and the associated contacts. The usage of the submodel template is slightly adjusted to fit the needs for the security risk assessment processes in accordance to the IEC 62443-3-2 standard. The third published submodel[29] "Hierarchical Structures enabling BOM" represents the information about the composition of entities within the SUC. This dissertation aligns towards the proposed structure from the submodel, but uses them in a decentralised approach.

9.4. Intermediate Summary

The fourth method step of information access transforms the previously defined information and logic for the automation of security risk assessments into practically usable models for the AAS as an industrial implementation of the DT. This addresses

[27] https://industrialdigitaltwin.org/en/wp-content/uploads/sites/2/2022/10/IDTA-02006-2-0_Submodel_Digital-Nameplate.pdf
[28] https://industrialdigitaltwin.org/en/wp-content/uploads/sites/2/2022/10/IDTA-02002-1-0_Submodel_ContactInformation.pdf
[29] https://industrialdigitaltwin.org/en/wp-content/uploads/sites/2/2023/04/IDTA-02011-1-0_Submodel_HierarchicalStructuresEnablingBoM.pdf

the deficit D4 of high abstraction levels from Section 4.3.3. First, the associated requirements regarding the needed information access for the automation of security risk assessments are analysed. In addition, this chapter offers an overview of the available approaches to make use of the DT concept within the OT domain and focuses on the AAS, OPC UA, AML, and MTP. Finally, all necessary information regarding the four phases of network segmentation, requirements & guarantees, risks, and attestation are integrated and modelled within the AAS on hierarchical levels that consist of three self-developed and three already published submodel templates. Figure 9.5 shows the overall dissertation structure and the current progress regarding the fourth method step of information access.

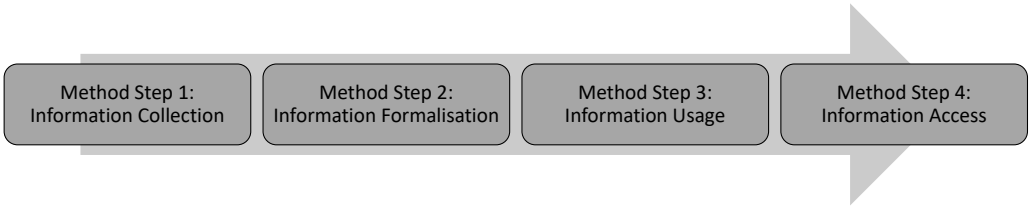

Figure 9.5: Dissertation progress after the fourth method step of information access

Part IV.

Result Discussion

10. Prototypical Implementation

This chapter contains the first part of the overall result discussion describing the prototypical implementation regarding the automation of the security risk assessment process. The idea is to elevate the specified concepts from the four previous method steps (Section 6 to Section 9) of information collection, formalisation, usage, and access into a practically usable expert system for automated security risk assessments. Thereby, it is possible to analyse, compare, and evaluate the automation degree of the proposed solution approach with the needed efforts and resources of the manual security risk assessments. Therefore, the next Section 10.1 provides the overview of the AAS tooling landscape including the different types AASs, the available programming frameworks to deploy AASs, and the methodology on how to retrieve and classify the required data to be filled into the AASs. Then, Section 10.2 illustrates the developed software architecture and programming conditions. In addition, the published open source GitHub repository is presented and explained as a basis for the subsequent evaluation.

10.1. AAS Tooling Landscape

The overall goal is to utilise the standardised AAS concepts as the chosen way of integrating the necessary information towards the automation of security risk assessments in a user-friendly and machine-readable manner (as discussed in Section 9.2). Therefore, the following three questions need to be answered within the following subsections:

1. Which type of AAS is most suitable for the described requirements?
2. Which software tools are available to be used for the AAS implementation?
3. How can the required input data for the AAS be retrieved and classified?

10.1.1. Available AAS Types

In general, the AAS metamodel specification offers various possibilities and different degrees of implementation styles, e.g. regarding communication capabilities, storage options, security features, or utilised protocols. This plethora of characteristics is also represented within the current state of available software tools for the AAS usage. Therefore, certain design decisions need to be made first to identify the most adequate software tool. First, the type of AAS needs to be determined. A

comparison of the three available types is provided within Table 10.1 [207, 210, 214, 220, 223]:

Table 10.1: Characteristics and comparison of the three available AAS types

Type	**Data Storage**	**Communication**	**Maturity**
Passive	(AASX) Files, e.g. JSON or XML	File transfer	Usable
Re-active	AASX files on a server	API via, e.g. HTTP REST	Usable
Pro-active	AASX files on a server	Autonomously peer-to-peer	Conceptual

This dissertation uses passive AASs stored in the text-based JSON format as a widely distributed and accepted way to model information. The JSON format is independent from any programming language and offers a machine- and human-readable style with low complexity, resulting in the lowest common denominator for information modelling and an easy handling of the files. It is equipped with clear simplicity of the format, universality in the usage on various platforms, and is very lightweight. In addition, the JSON files can be integrated into the official AASX[1] package file format for AASs which generally works similar to a zipped file containing all needed information and offering additional functionalities, e.g. signing, for an AAS. This enables an I4.0- and AAS-compliant data representation and encapsulation [220]. The created JSON files containing the AASs are integrated into the security risk assessment process implementation due to the advanced degree of maturity, available software tools and libraries, low storage requirements, and the adequate exchange capabilities. An upgrade towards the re-active or the pro-active AAS type might be beneficial to increase the autonomous communication capabilities, but for the time of writing this dissertation it is regarded as future work. In addition, up to now no advanced communication capabilities are necessary. The current version of the prototypical implementation directly reads the AAS files and processes the contained information within the algorithmic logic. Furthermore, the data flow is just unidirectional (Read) from the AAS files into the algorithm, a bidirectional (Read & Write) communication, e.g. to write the assessment results or updated values to the AAS as presented in [220], is not needed at the moment and therefore not implemented right now.

10.1.2. Software Tool Overview

The landscape of available software tools for an AAS implementation is diverse and heterogeneous with many different approaches and concepts available each covering certain aspects or characteristics [211]. Various software tools are currently under development, most of which are not interoperable, represent different versions of the AAS metamodel specification, or provide different communication capabilities [211].

[1] https://industrialdigitaltwin.org/wp-content/uploads/2023/04/IDTA-01005-3-0_SpecificationAssetAdministrationShell_Part5_AASXPackageFileFormat.pdf

The Eclipse Foundation[2] and the IDTA[3] can be regarded as central information hubs. Furthermore, additional surveys and summaries about the AAS and the associated software tools can be found within a plethora of related research publications [207, 210–212, 220, 224–226]. At the time of writing this dissertation, ten software tools in total are known. Six of the minor software tools are mentioned in the following for the sake of completeness, but are not further regarded due to the mentioned reasons:

1. Eclipse AAS Web Client (AASWC)[4] → Pure GUI for displaying AASs hosted within the AASX Server
2. Eclipse AAS Model for Java (AAS4J)[5] → Solely a library for translating the AAS metamodel into Java classes and XML, JSON, or RDF files
3. AAS by the German Research Center for Artificial Intelligence (DFKI)[6] → An AAS extension with very limited features based on the well-established Eclipse BaSyx implementation [211]
4. AAS Azure Services[7] → Inadequate documentation and currently no running code [211]
5. CoreAAS[8] → Outdated and no longer maintained OPC UA implementation in TypeScript [225]
6. i40-aas[9] → Already archived and integrated into other projects by the IDTA due to high complexity and missing documentation [224]

This results in four software tools left over for a more detailed investigation and analysis if adequate for the requirements of this dissertation [211]:

1. AASX Server + AASX Package Explorer
2. Eclipse BaSyx
3. FA³ST
4. NOVAAS

[2]https://projects.eclipse.org/projects/dt/governance
[3]https://admin-shell-io.com
[4]https://projects.eclipse.org/projects/dt.aaswc
[5]https://projects.eclipse.org/projects/dt.aas4j/developer
[6]https://github.com/dfkibasys/asset-administration-shell
[7]https://github.com/JMayrbaeurl/opendigitaltwins-aas-azureservices
[8]https://github.com/OPCUAUniCT/coreAAS
[9]https://github.com/SAP-archive/i40-aas

The IDTA hosts the **AASX Server** and the **AASX Package Explorer** software tools on GitHub[10], originally coming from an Eclipse project[11]. As all other software tools, they are still under development currently and subject to changes, especially catching up with the currently ongoing AAS metamodel specification updates. The provided framework consisting of both software tools offers two main functions:

1. Visually creating and viewing AASs via the AASX Package Explorer
2. Providing an interface towards file-based AASs via the AASX Server

Both software tools are platform-neutral with corresponding Docker images that are available for personal computers, are implemented using C# and the .Net framework, support the AASX file format, and enable communication via natively supported APIs, e.g. using HTTP REST, MQTT, or OPC UA [224]. Furthermore, essential security features are given, e.g. role-based authentication and token-based authorisation to access AAS contents and the usage of Transport Layer Security (TLS) for the communication protocols [211, 212, 220].

The **AASX Server** is the currently most prominent software tool for modelling and managing AASs by providing adequate file storage, a defined interface, and suitable communication possibilities [211]. It is distributed under the Apache v2.0 license. Nevertheless, it is still under development by the IDTA and is subject to constant changes and updates. In addition, at the moment it seems to be outdated and not supporting the latest AAS metamodel and API specifications, resulting in low synchronisation capabilities between the physical component and the AAS [211]. The AASX Server itself offers a command line interface and a basic GUI for the configuration and maintenance of the software tool. Generally, the AASX Server is available in three different variants [211]:

1. core → Only the server with a command line interface
2. blazor → Enhancing the server with a basic GUI
3. windows → Running on Windows without administrator privileges

The **AASX Package Explorer** enhances the server functionality with an extensive GUI to create, load, change, and store AASs in the AASX file, JSON, or XML formats via an editor [211]. Therefore, it enables also less technically-inclined users[12] to experiment with the AAS concepts and accelerates the development and distribution of the AAS, especially when starting with new projects from scratch [224]. The open source software tool can be found publicly on GitHub[13] under the Apache v2.0 license.

[10]https://github.com/admin-shell-io/aasx-package-explorer
[11]https://github.com/eclipse-aaspe/aaspe
[12]https://projects.eclipse.org/projects/dt.aaspe/reviews/creation-review
[13]https://github.com/admin-shell-io/aasx-package-explorer

BaSyx is a well-established and widely-used software tool suite, containing e.g. server, registry to manage the AAS organisation, web client, asset synchronisation, and visualisation, to cover the AAS lifecycle as a middleware in a service-oriented manner [226]. It is now hosted by the Eclipse Foundation[14], originating from the BaSys 4.0[15] and the follow-up BaSys 4.2[16] research projects funded by the German Federal Ministry of Education and Research (BMBF) [212, 226]. The code is open source and publicly hosted on GitHub[17] under the MIT license in a platform neutral manner in the form of Docker images [211]. The supported programming languages are Java, C++, Rust, .Net, and Python (formerly being PyI40AAS[18]) [211, 224]. The software tool offers AASX file support using the XML and JSON file formats and for type 2 re-active AASs the protocols HTTP REST, MQTT, and OPC UA for a bidirectional communication [226], making it partly compatible with the AASX Package Explorer as well [224]. Furthermore, Eclipse BaSyx provides a rule-based security approach to configure access rights on a resource level and enables rudimentary sanity checks of the modelled information [211].

The Fraunhofer Advanced AAS Tools for Digital Twins (**FA³ST**)[19] are publicly stored on GitHub[20] under the open source Apache 2.0 license. They include approaches to model, create, and use DTs following the AAS specification, currently implementing the version 3 of the AAS metamodel [212]. The software tool implements a type 2 re-active AAS, focuses on the asset synchronisation, and provides an easy usage for non-experts with possibilities to develop own tools for the AAS [211]. The FA³ST approach is still under development and lacks major tooling support at the moment, such as client libraries, registries, visualisation, and the integration of security-related features [211].

The NOVA School of Science and Technology develops the NOVA Asset Administration Shell **(NOVAAS)** within the Horizon 2020 PROPHESY[21] EU research project. The source code is publicly available on GitLab[22] and is published under the EUPL v1.2 license [211]. In general, NOVAAS follows a no/low-code flow-based programming concept using Node-RED (based on JavaScript) and offers an AAS metamodel implementation using AASX files in the JSON format, a communication API via HTTP REST, and a GUI to create dashboards [211, 224]. The overall framework is platform-neutral with a Docker image available, but has only a limited amount of functionality, currently lacking, e.g. a client library, a model editor, model validation, or a registry [224]. Nevertheless, NOVAAS offers a partial compatibility towards other software tools, e.g. the AASX Server & AASX Package Explorer, offers

[14] https://projects.eclipse.org/projects/dt.basyx
[15] https://www.basys40.de
[16] https://www.eclipse.org/research/projects/basys_42
[17] https://github.com/eclipse-basyx
[18] https://git.rwth-aachen.de/acplt/pyi40aas
[19] https://www.iosb.fraunhofer.de/en/projects-and-products/faaast-tools-digital-twins-asset-administration-shell-industrie40.html
[20] https://github.com/FraunhoferIOSB/FAAAST-Service
[21] https://prophesy.eu
[22] https://gitlab.com/novaas/catalog/nova-school-of-science-and-technology/novaas

basic visualisation via the browser, and has basic security features implemented, such as role-based authentication and authorisation [211, 224].

The following Table 10.2 summarises the characteristics of the four presented main software tools and illustrates their advantages, disadvantages, and differences among each other. The analysis is based on the references [211, 212, 224–226] and is used to decide about the usage of an AAS software tool for this dissertation. The defined criteria are answered for each solution by distinctive facts, e.g. regarding the licensing model, and a three-tier evaluation has been done based on the following schema: "✔" indicates a suitable maturity of the respective functionality, "∼" represents a moderate maturity of the respective functionality, and "✗" indicates a missing functionality. In summary, FA³ST and NOVAAS are only marginally-used software tools offering specific and limited functionality for the implementation of AASs and are therefore not further regarded. Eclipse BaSyx and the AASX Server + AASX Package Explorer are the most prominent software tools at the moment mainly used within the industrial automation research domain and the associated projects. Both frameworks offer a similar range of functionality with comparable characteristics and functionalities: Open source licensing model, publicly available source code, adequate support of the AAS metamodel specification, AASX file integration, and various communication protocols for the API.

In the end, the advanced usability, the extensive GUI, the support of the widely active IDTA, and the intuitive approach of the AASX Server + AASX Package Explorer are weighted higher and concluded to the decision to use this software tools for the integration into the prototypical implementation of this dissertation. In this way it is possible to create, adapt, and use AASs in an easy and adequate manner to provide the information access for the automated security risk assessment process.

Table 10.2: Comparison of the four presented software tools for AAS implementation: AASX Server + AASX Package Explorer, Eclipse BaSyx, FA³ST, and NOVAAS

Criteria	**AASX Server + AASX Package Explorer**	**Eclipse BaSyx**	**FA³ST**	**NOVAAS**
Stakeholder	IDTA	Research project	Fraunhofer	Research project
License	Apache 2.0	Apache 2.0	MIT	EUPL 1.2
Programming language	.NET	Java, C++, Rust, .Net & Python	Java	Node RED
Metamodel version	2.0.1 (v3 is undocumented)	2.0.1	3.0RC01	2.0.1
Model formats	JSON & XML	JSON & XML	JSON, XML & RDF	JSON
APIs	HTTP, MQTT & OPC UA	HTTP, MQTT & OPC UA	HTTP & OPC UA	HTTP
GUI	✔	∼	∼	∼
Security	✔	✔	✗	∼
Usability	✔	∼	∼	∼

The next subsection covers the topic of data retrieval and classification to ensure that the correct information are available within the implemented AAS.

10.1.3. Data Retrieval and Classification

So far, it has been discussed and described which kind of information are needed for the automation of security risk assessments as well as how and where they are going to be stored. What is still missing is the classification of these information and when, from where, and by whom they are going to be retrieved into the AASs in real-world scenarios requiring a high practicability for industrial use cases. The beforehand mentioned method step of information access (see Section 9) summarises the integration of the AASs into the security risk assessment algorithm. In general, there are various types of data sources available, e.g. analogue or digital [120], and the associated topic of knowledge and asset management opens up an additional research field of its own [44]. It is important to cover the complete security knowledge lifecycle consisting of determination, acquisition and creation, maintenance, usage, evaluation, and exclusion of information (as defined within [180]). Therefore, this subsection illustrates the Current Mode of Operation (CMO) for data retrieval and classification used within this dissertation and describes the outlook towards a possible Future Mode of Operation (FMO). Nevertheless, the complete handling and coverage of this research topic is out of scope for this dissertation and only mentioned here for the sake of completeness. The current state of the developed prototypical implementation inherits the following boundary conditions:

- The algorithm for the security risk assessment is prepared and available
- Definition of a digital representation of the SUC as AASs
- Information about network topology, asset characteristics, communication interfaces, and preceding safety assessments are present
- Connectivity towards external databases for, e.g. technical vulnerabilities or human adversary characteristics, is given

The CMO results in an increased degree of automation for security risk assessments, but still requires manual preparation for the necessary information availability via AASs, data retrieval, and classification (as described for the example of the SL-T determination [94] from Annex B in Section 15). The following Table 10.3 summarises the necessary data input that has to be prepared and provided for the defined security risk assessment process as shown within Section 6 via the swimlane methodology (also see Annex A in Section 14). It is described what kind of information (data name & domain) is required, where it needs to come from (source & stakeholder), how it can be characterised (specificity & format), and how often it changes (alteration). The data name is a free text and describes the collected information from Section 6. In [44], typical domains of origins for security-related information are categorised, such as asset or system level and external influences. The source classification is adapted from [120] and describes the way of data retrieval, e.g. via manual extraction from documents or tool-based. As described before (see Section 3.1), the main industrial stakeholders of component manufacturer, system integrator, and asset owner can

be each assigned to information which they can provide. In [120], the specificity of information is presented, being either system-dependant or system-independent. The format of the required information can be compared to a variable type representing the storage option in an algorithmic environment. At last, the alteration categorised within [44] describes the change rate of information ranging from never via on demand to hourly.

Table 10.3: Overview of the required information retrieval and classification for the automation of security risk assessment processes

Data Name	**Domain [44]**	**Source [120]**	**Stakeholder**	**Specificity [120]**	**Format**	**Alteration [44]**
Asset inventory	System	Tool	Asset owner	System-dependent	Database	On demand
Asset ID	Asset	Documentation / Tool	Component manufacturer	System-dependent	String	Never
Safety suitability	Asset	Documentation	Component manufacturer	System-dependent	Boolean	Never
Module ID	System	Documentation	Asset owner	System-dependent	String	Never
Responsible	Human entity	Documentation	Asset owner	System-dependent	String	On demand
Accountable	Human entity	Documentation	Asset owner	System-dependent	String	On demand
Physical ports	Asset	Documentation / Tool	Component manufacturer	System-dependent	String	On demand
Threats	External influences	Public database	Security community	System-independent	String	Hourly
Attacker skills	External influences	Public database	Security community	System-independent	String	Annually
Attacker resources	External influences	Public database	Security community	System-independent	String	Annually
Countermeasures	Asset / System	Documentation	Asset owner	System-dependent	String	On demand
SL-C	Asset	Documentation	Component manufacturer	System-independent	Integer	Never
SL-A	Asset	Documentation	System integrator	System-dependent	Integer	On demand
SL-T	Asset / System	Risk assessment	Asset owner	System-dependent	Integer	Annually
Vulnerabilities	External influences	Public database	Security community	System-independent	String	Hourly

To enhance the presented overview of necessary data for the security risk assessment process, a more detailed example is provided for better comprehensibility, namely the asset characteristic of safety suitability from Table 10.3. This data input is required for the network segmentation of the SUC into zones and conduits following the IEC 62443 standard based on the safety characteristics of the assets. The required information from the assets can be described semantically using ECLASS[23] as an inter-sectoral classification system based on international standards describing products, materials, or services and their properties [200]. The alignment towards this already accepted and widely-used formalisations enhances the credibility and reusability of the approach. In addition, the following characteristics of the required data input are defined to improve the comprehensibility of the underlying database for the security risk assessment process and the data classification respectively retrieval methods:

- Name: SuitableForSafetyFunctions
- Representation [57]: ECLASS
- Requirement source [120]: Standard / Guideline
- Reference: IEC 62443-3-2 ZCR 3.1 & ZCR 6.4 e)
- Purpose [56]: Network segmentation

[23] https://eclass.eu/en/eclass-standard/search-content

- Identification [57]: IRDI → 0173-1#02-BAD722#009
- Data source: Component manufacturer
- Data creation: Manually during engineering
- Knowledge type [44]: Class
- Articulability: Direct
- Ordinality: Qualitative
- Consistency: Discrete
- Perception: Objective
- Storage: MiscComponent AAS submodel (from Section 9.3)

The presented CMO to retrieve and classify the necessary information for the utilisation with AASs within this subsection still has to be performed manually. Nevertheless, once the information are prepared, the automated security risk assessment process can be performed and used to generate results adequate for the corresponding operator. The additionally required research on the topic of data retrieval and classification is out of scope for this dissertation. Associated topics and related work for a possible FMO are summarised here:

- Available tools for the automated data retrieval and classification (see Section 4.1)
- Consistent knowledge management for the complete asset lifecycle [44]
- Integration of (legacy) approaches for asset identification, e.g. Common Platform Enumeration (CPE)[24] or Security Content Automation Protocol (SCAP)[25]
- Alignment towards upcoming approaches from the security domain, e.g. Common Security Advisory Framework (CSAF)[26] or SBOM[27] with VEX

The subsequent section illustrates the usage of the presented and evaluated software tools for the prototypical implementation and the overall software architecture as the basic foundation of the automated process for security risk assessment processes.

[24]https://nvd.nist.gov/products/cpe
[25]https://csrc.nist.gov/projects/security-content-automation-protocol
[26]https://oasis-open.github.io/csaf-documentation
[27]https://www.cisa.gov/sbom

10.2. Software Architecture

10.2.1. General Overview

The developed prototypical implementation of this dissertation is based on the swimlane modelling from Section 6 and is programmed with several software frameworks and libraries. In general, a TRL of 6 is achieved representing the development stage of a technique which is demonstrated on an actual system within a real-world scenario [2]. This results in a LOA of 4 expressing a system which is in full control of certain situations and the human operator is just there to supervise the results (following the definitions within Section 4.1) and in a LOK of 1 reflecting a requirement for a typical analyst with simple means having minimum resources and basic security-related skills (following the definitions within Section 5.1). It can be found publicly in an open source GitHub repository[28] containing a dedicated user guidance on how to get started, test, and extend the developed prototypical implementation. This enables also external and inexperienced users to utilise the presented exemplary test cases based on an automated security risk assessment of the Customisable Production System demonstrator in the SmartFactoryOWL[29] for their own purposes. In addition, the repository includes various other contents:

- *doc* → Additional documentation including figures of the used AASs, the overall process, and the provided test cases
- *src* → Program code in Python with comments of the prototypical implementation
- *knowledge* → Files with the specified and formalised expert knowledge necessary for the automated security risk assessment
- *aas_examples* → AASs acting as the basis for the three test cases, the submodel templates as AASX files, and a random SL generator to create own AASs with different values for the SLs

Figure 10.1 illustrates the general software architecture of the complete prototypical implementation. The overall environment can be divided into internal and external aspects indicated by the black dashed vertical line. The main part of this dissertation lies within the internal side of the implementation. There, the core of the security risk assessment process as the algorithm (represented in orange) is realised in the Python programming language. Therefore, the software can be run everywhere with an adequate version of the Python environment installed, e.g. personal computers or laptops, mobile devices, servers, cloud services, and even low-resource hardware. It contains the implementation of the methodology conceptualised and developed within the previous sections. In addition, specifically created information sources are read by

[28] www.github.com/auto-s2/security-risk-assessment
[29] https://smartfactory-owl.de/?lang=en

the algorithm. This includes the exemplary AASs following the metamodel version from April 2023[30] (shown in blue) to represent the Customisable Production System demonstrator from the SmartFactoryOWL and the manually created knowledge base in cooperation with the AutoS² research project (shown in purple). The results of the security risk assessment process are communicated via the command line interface and are saved to the attestation document within the Portable Document Format (PDF) file format (shown in green). The external aspects of the prototypical implementation contain two additional sources (shown in yellow): Technical information about vulnerabilities based on CVE[31] and CVSS v3.1[32] from the NVD[33] and attacker characteristics from the MITRE ATT&CK ICS framework v13.1[34] as security expert knowledge. Possible updates to CVSS v4.0[35] and MITRE ATT&CK ICS v15[36] are currently suspended due to the lack of fitting information for the required knowledge base of this dissertation and are therefore part of the future work. This overall architecture also fulfils the typical expert system requirement regarding a separation into knowledge base and processing (as defined within Section 4.2) [44].

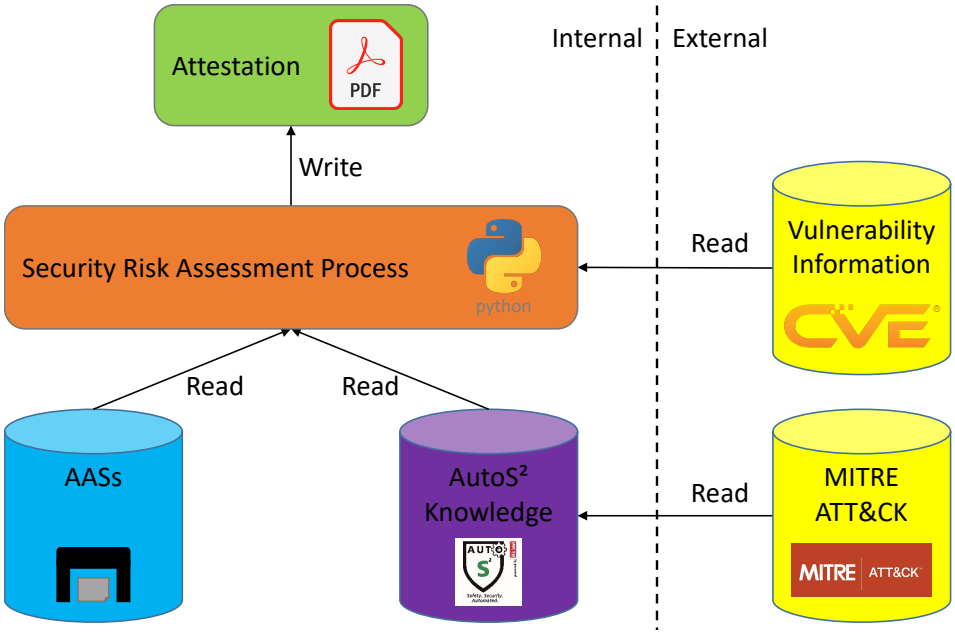

Figure 10.1: Software architecture of the prototypical implementation

[30] https://industrialdigitaltwin.org/en/wp-content/uploads/sites/2/2023/06/IDTA-01001-3-0_SpecificationAssetAdministrationShell_Part1_Metamodel.pdf
[31] https://www.cve.org
[32] https://www.first.org/cvss/v3.1/specification-document
[33] https://nvd.nist.gov/vuln/search
[34] https://attack.mitre.org/versions/v13
[35] https://www.first.org/cvss/v4.0/specification-document
[36] https://attack.mitre.org/versions/v15

The required Python modules are collected and stated within the *requirements.txt* file[37] as a collection of software dependencies. Other than the displayed modules here, only standard Python modules are used, e.g. *time*, *os*, or *json*:

- *requests*[38] version 2.25.1 → HTTP library for the retrieval of vulnerability information from the NVD based on CVE and CVSS
- *openpyxl*[39] version 3.0.9 → Python module to read and write Excel files for integrating the knowledge base into the program code
- *fpdf*[40] version 1.7.2 → PDF generator used for the creation and storage of the attestation document

The overall structure of the prototypical implementation can be abstractly described by the following flowchart depicted within Figure 10.2. It represents the four automated phases of the security risk assessment process (network segmentation, requirements & guarantees, risks, and attestation) on a general level and shows the associated inputs and outputs necessary for the algorithm. A more detailed description of each function block of the flowchart is already provided within the original specification of the swimlane model in Section 6.2 of the first method step of information collection. Therefore, these contents are not repeated here.

In general, the prototypical implementation is based on the experiences gained and the concepts developed from the three manually performed security risk assessments described within Section 5.1. One of these inspected and assessed demonstrators from the SmartFactoryOWL in Lemgo is the Customisable Production System. This system was chosen as the exemplary and representative SUC for this dissertation and the resulting automation of security risk assessment processes due to the availability of information (already partly in the form of AASs) and the responsible stakeholders, the medium level of complexity and size, the variety of components and manufacturers, and the real-world impact of safety-related issues regarding the integrated laser engraving machine. Figure 10.3 shows the general network topology of the Customisable Production System including the components, their physical connections for the internal and external communication interfaces (black lines), and their affiliation within the modules of the SUC (light grey boxes). In addition, the figure also illustrates the network segmentation into zones and conduits based on the concepts from the IEC 62443 standard, which is one of the automated results from the prototypical implementation. The Customisable Production System consists of three modules (Laser, Conveyor, and Cabinet). These modules consist of components that are safety-relevant (yellow boxes) and components that are not safety-relevant (green boxes). The demonstrator itself is configured within one local subnet and the connection to the public network (light blue cloud) is enabled via the core switches

[37] https://github.com/auto-s2/security-risk-assessment/blob/main/src/requirements.txt
[38] https://pypi.org/project/requests
[39] https://pypi.org/project/openpyxl
[40] http://www.fpdf.org

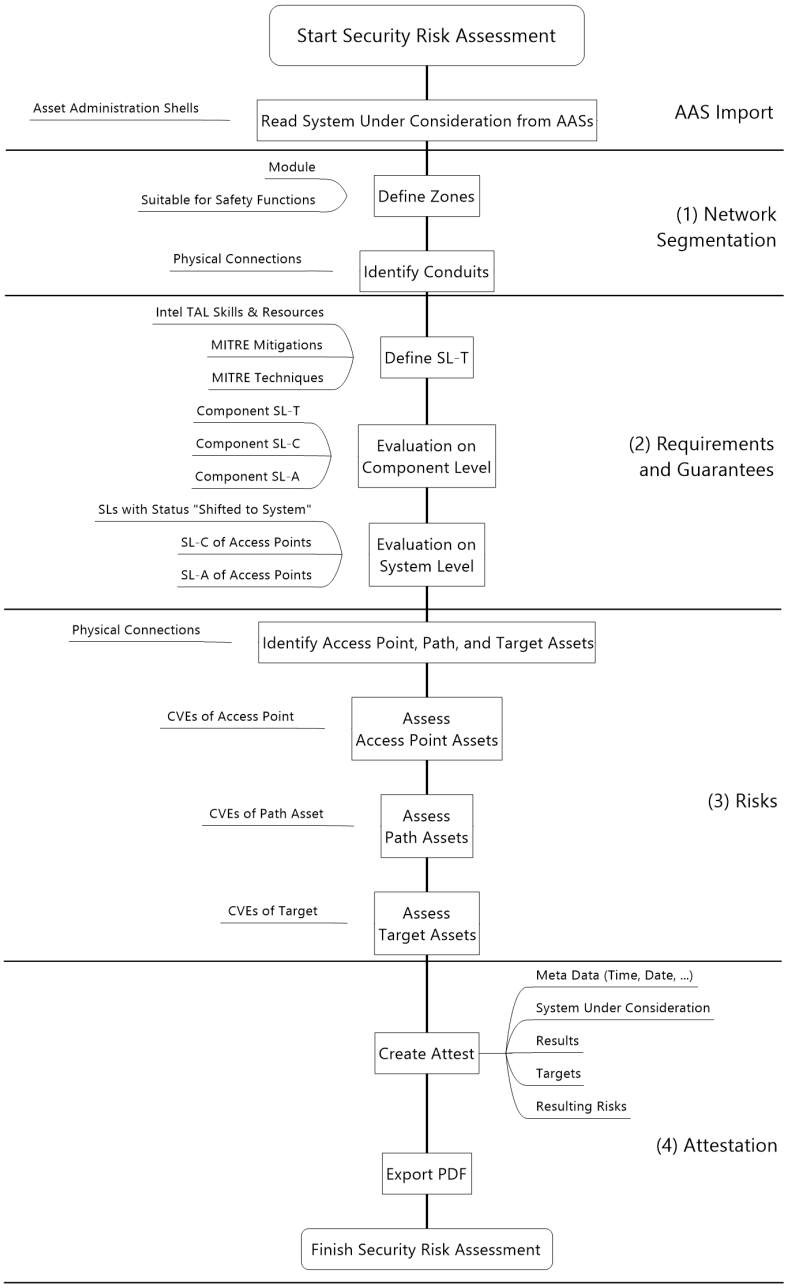

Figure 10.2: High-level process overview of the implemented security risk assessment

and the central firewall of the SmartFactoryOWL with a standard rule set denying all incoming traffic and allowing all outbound traffic by default and without any specific port configurations. Furthermore, the Cabinet module was artificially enhanced by three additional safety-related PLCs to improve the use case variability and to create the possibility to test different scenarios. These additional PLCs are considered to be

bridged for the communication of safety-relevant protocols. Based on the respective component characteristics, the components are grouped into zones (dark grey boxes). This representative system architecture and the associated network topology are used for the further evaluation of the prototypical implementation of this dissertation.

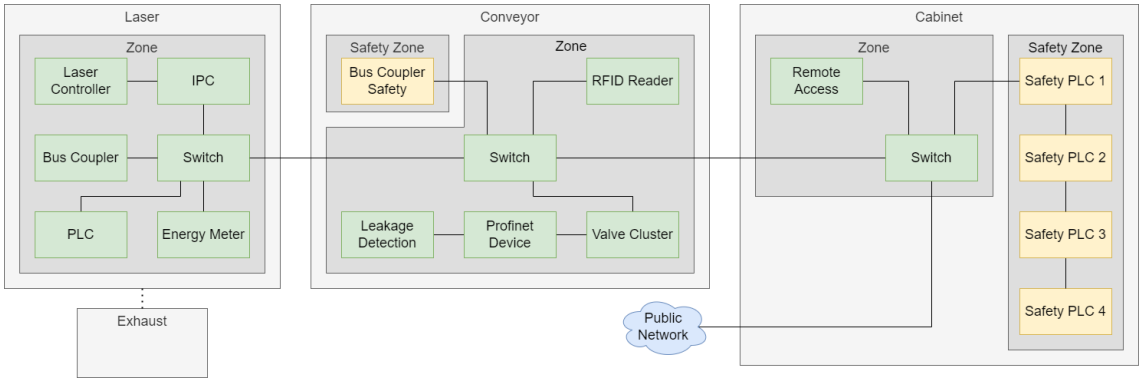

Figure 10.3: Network topology of the Customisable Production System from the SmartFactoryOWL

10.2.2. Basis for Evaluation

To test and evaluate the prototypical implementation, three test cases are developed and the modelled information are represented via distinctive AASs. These three exemplary test cases differ in technical vulnerabilities represented by varying CVE and CVSS values of the safety-relevant PLCs within the cabinet module which are stored within the knowledge section of the GitHub repository[41]. This mixed presence of technical vulnerabilities can be typically found due to varying device series or different patch states of the assets. In the first example, all four safety-related PLCs have technical vulnerabilities described CVE and CVSS values leading to resulting risks and enabling an attacker to move further within the network via lateral movement. Within the second example, the safety-related PLC 3 has no vulnerabilities. As a result, the safety-related PLC 3 is protected and is blocking the path for the lateral movement of possible attackers also securing the trailing components. The third and last example contains no technical vulnerabilities for the safety-related PLCs. As a result, there are no resulting risks. Figure 10.4 illustrates the three described examples in a graphical way.

The test case "Example 2" will be taken as the further example for the description here and for the evaluation to describe the results of the automated security risk assessment process within Section 11. Figure 10.5 shows a copy of the created attestation document and the included contents as an abstracted summary of the security-relevant results for the operator. This is highly important because the SUC needs to be seen as a socio-technical system and the human factor is still critical

[41] https://github.com/auto-s2/security-risk-assessment/tree/main/knowledge

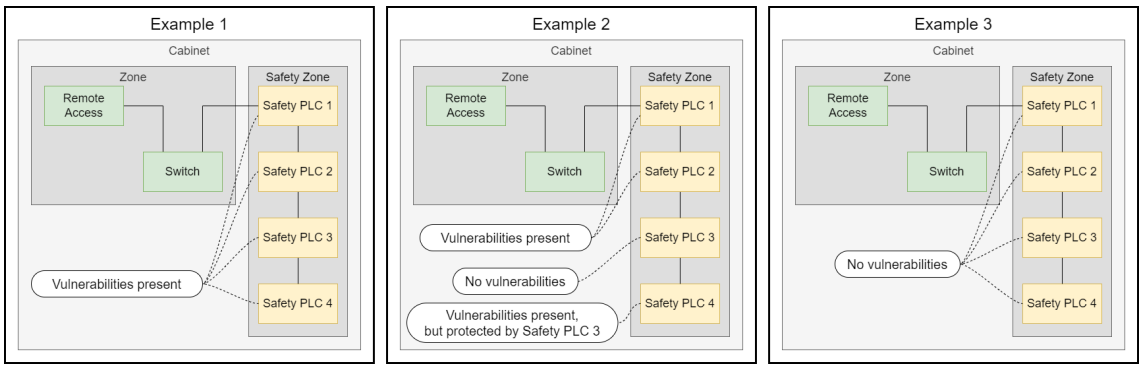

Figure 10.4: Overview of the three implemented example test cases of the cabinet module in the Customisable Production System based on different technical vulnerabilities with varying CVE and CVSS values

for the final decision making [37]. In addition, the results can also be displayed in the command line of the device executing the prototypical implementation. The attestation document is separated into four main parts:

1. Information
2. SUC
3. Results
4. All Targets and Resulting Risks

The first part contains general information and meta data about the performed security risk assessment process itself: Date and time, a unique identifier, the computing time, the username of the operator, and the concurred errors during the runtime of the algorithm. These errors can also be found within the attestation on the following pages of the PDF document to gain insights into more details. The second part covers the information about the SUC from a network topology point of view enhanced with security-related facts and recommendations for further actions adapted from the IEC 62443 standard, such as zones, conduits, CRs, or SRs. The following part includes a summary by accumulating the various results from the different components regarding the resulting risks and the associated lowest attacker skills respectively resources adapted from the Intel TAL which are applicable to unmitigated security risks. The fourth and last part displays all target assets (as defined within Section 6.2.4) which are in scope of the security risk assessment process and their resulting security risks based on the present and known technical vulnerabilities and the investigated attacker characteristics and techniques. In the presented second example five target assets are identified and assessed. As always with regard to security, a 100% guarantee of security risks being absent cannot be given and the security risk assessment results are valid at the point of creation based on the then available information. Two of them have a resulting risk of *"VERY HIGH"* and

the remaining three components have no identified risks which is indicated by *"NO RELEVANT RISK"*. In the end, the information from the attestation document can be used by the operator to decide about a possible risk treatment (as defined within the ISO 27005 and the ISO 31000 standards) of accepting, transferring, mitigating, or terminating the risk.

Attestation

Information

Date and Time of Attestation:	01.05.2024 09:45:20
Attestation ID:	0x13c947b8fb2020e
Algorithm Computing Time:	07.42 Seconds
Operator (Username):	CPS_Security
Errors from the Algorithm:	0

System under Consideration

Total Number of Modules:	4
Total Number of Components:	18
Total Number of Zones:	5
Total Number of Targets:	5
Total Number of Unmitigated SRs:	79 in 5 Zones
Total Number of Reconf. Adv. SRs:	16 in 3 Zones
Total Number of Reconf. Adv. CRs:	351 in 18 Components

Results

Highest Risk of all Target Assets:	VERY HIGH
Lowest Intel TAL Attacker Skill (unmitigated):	Adept
Lowest Intel TAL Attacker Resource (unmitigated):	Organization

All Targets and Resulting Risks

CabinetSafetyPLC1	VERY HIGH
CabinetSafetyPLC2	VERY HIGH
CabinetSafetyPLC3	NO RELEVANT RISK
CabinetSafetyPLC4	NO RELEVANT RISK
ConveyorSafetyBusCoupler	NO RELEVANT RISK

Figure 10.5: Attestation results from the "Example 2" of the security risk assessment process in the form of a PDF document

11. Evaluation

This section contains the second part of the overall result discussion based on the actual evaluation after the description of the prototypical implementation and builds upon previous works by the author of this dissertation and extends the published results from [179, 227] and the planned publication for the ETFA 2024 which is still under review during the time of writing. In general, an evaluation based on real-world technical systems from the OT domain, especially with regard to the topic of security, is challenging as characteristics are typically only qualitative and subjective [196]. In most cases, a trade-off between accuracy, completeness, and cost of the evaluation process is necessary which determines how detailed the evaluation can be done. This is also dependant on the type of evaluation for which, generally, four main methods are available [228]:

- Analysis: A mathematical approximation of the given system is constructed and equations are derived to describe the characteristics.
- Emulation: Mimicking the outer behaviour of the investigated system. The internal states do not have to accurately reflect the reality.
- Simulation: Modelling of the underlying states of the target system including a correct representation of the internal states.
- Experiments: Real examples of hardware and software components are used to build up a prototypical implementation for measurements.

The further evaluation of this dissertation is based on the presented prototypical implementation (from Section 10) as an **experiment** representing a typical SUC as defined within Section 1.2 covering the application scenario. This procedure will show the applicability of the results and provides the possibility for a comprehensive evaluation of the expert system for automated security risk assessments [137]. The overall structure of the evaluation is based on the following three parts [196]:

1. Verification: Internal review to determine whether the proposed solution approach satisfies the requirements and if the system was built correctly.
2. Validation: External review to determine whether the proposed solution approach complies with the functional specification and if the system is correct.
3. Credibility: Security-related expert systems need to have high trustworthiness, as incorrect models will result in inadequate security decisions.

The following sections will cover the three main aspects of the overall evaluation. The verification is covered in Section 11.1 including the comparison and evaluation of general requirements and the respective requirements from the four main method steps of information collection, formalisation, usage, and access. The validation is shown in Section 11.2 and is based on the definition of a reference set of data from a human domain expert for the comparison of manual and automated security risk assessment. This part is complemented with the final check of the hypotheses from the problem statement in Section 1.3. The last part covers the interpretation of the overall outcomes of this dissertation to achieve a high degree of credibility. This is also presented in Section 11.3 regarding the evaluation of the generalisability and the concluded recommendations of this dissertation.

11.1. Verification: Requirements Analyses

The verification is summarised within the upcoming subsections 11.1.1 to 11.1.5, each presenting the associated requirements, a comparison of the most relevant publications from the related work (see Section 4.3 for respective descriptions, characteristics, advantages, and disadvantages), and a short summary. The five most relevant publication series of related work are summarised and enumerated for the rest of this section as follows (taken from Table 4.2 in Section 4.3.2):

- **#1)** M. Eckhart: Tool for the automated risk identification based on logic rules and AML [146, 175]
- **#2)** C. Tebbe: Consistent OT security knowledge management for the lifecycle of manufacturing systems [44]
- **#3)** S. Fluchs & E. Tastan (NAMUR WG 1.3): Reference model for security engineering [32, 47, 128, 182–184]
- **#4)** B. Brenner: Network-based IDS for OT environments being capable of coupling safety and security [185]
- **#5)** C. Tebbe: Partly automated security analysis approach based on a rule- and knowledge-based system [46]

The expert system for automated security risk assessments from this dissertation is named and abbreviated as SRA4IACS from here on (as an improvement of the tool already presented in [229] in 2020). Furthermore, the identified challenges (see Section 3.5) and deficits (see Section 4.3.3) are addressed by the four method steps of information collection, formalisation, usage, and access (from Section 6.1 to Section 9.1). The requirements analyses are assessed in a three-tier style to clarify the subjective and qualitative characteristics based on a text-based explanation (in alignment to the proposal from [74]). Each tier of evaluation is enhanced with a numerical value of points to also enable a quantitative summary based on the distinctive scores at the end.

- ✔ → Fulfilment of a requirement → 1 Point
- ∼ → Partial or insufficient fulfilment of a requirement → 0.5 Points
- ✗ → Non-fulfilment of a requirement → 0 Points

The respective details of the requirements analysis can be found within the following subsections 11.1.1 to 11.1.5. The overall results of the verification are then summarised within Section 11.1.6 afterwards.

11.1.1. General Characteristics

Firstly, the fact sheets of the general requirements regarding the prototypical implementation are presented here (all adapted from statements by [74]) as an addition to the requirements from the four method steps. Each general (abbreviated with "G") requirement (abbreviated with "R") can be identified via its short name as the title in bold and a unique identifier in the following style "R-G-#". Furthermore, a description of the requirement is given to summarise the demanded contents. The three different verbs which are used, represent the relevance of each requirement in a prioritising order (in accordance to the RFC 2119[1] and the ISO/IEC Directives[2]): High = *"shall"* / Medium = *"should"* / Low = *"may"*. All requirements are not directly measurable and are therefore evaluated in a qualitative manner based on a subjective justification by the author of this dissertation later on.

Requirement: Automation

- ID: R-G-1
- Description: The security risk assessment *shall* be implemented as a software tool to reach an adequate LOA and TRL for a high degree of automation.

Requirement: Openness

- ID: R-G-2
- Description: The automated security risk assessment process *shall* be available as open source to interested researchers and other industrial stakeholders.

Requirement: Usability

- ID: R-G-3
- Description: The abstraction and presentation of the automated security risk assessment results *should* be suitable for asset owners and system integrators with their associated LOK via a dedicated GUI.

[1] https://datatracker.ietf.org/doc/html/rfc2119
[2] https://www.iso.org/sites/directives/current/part2/index.xhtml

Requirement: Scalability

- ID: R-G-4
- Description: The automated security risk assessment *may* be able to process varying sizes of IACSs based on reasonable efforts and performance.

Requirement: Evidence

- ID: R-G-5
- Description: The automated security risk assessment *may* be based on specific configuration information of the SUC.

Secondly, Table 11.1 evaluates the five selected most relevant publication series from the related work (see Section 4.3.2) and compares them with the expert system from this dissertation based on the five defined general requirements.

Table 11.1: Evaluation results of the general requirements

ID	Requirement	#1	#2	#3	#4	#5	SRA4IACS
R-G-1	Automation	✔	✗	✗	✔	✔	✔
R-G-2	Openness	✔	∼	∼	✗	✗	✔
R-G-3	Usability	✗	✗	✗	✗	✔	∼
R-G-4	Scalability	✔	✗	✔	✗	✗	✗
R-G-5	Evidence	∼	∼	∼	✔	∼	✔
Evaluation Score:		**3.5**	**1.0**	**2.0**	**2.0**	**2.5**	**3.5**

R-G-1: The focus of this dissertation is set on the automation to support security risk assessment processes, especially for SMEs within the OT domain. This is achieved by the SRA4IACS approach as well as #1, #4, and #5. The works of #2 and #3 do not offer any automation due to the limited conceptual basis.

R-G-2: The complete software tool SRA4IACS is available in an open source manner, the same holds true for #1. The other related work only offers descriptive concepts (#2 and #3) or do not disclose any source code (#4 and #5).

R-G-3: To abstract and prepare results for the operators is a difficult requirement regarding usability. #5 offers an interactive GUI, whereas SRA4IACS solely provides security-related results summarised via a PDF file or via the command line, and the works of #1, #2, #3, and #4 do not offer any result overview in an abstracted manner reducing the necessary LOK.

R-G-4: The scalability is an important factor regarding the utilisation from small and simple systems up to large and complex ones. #1 offers various implementation aspects to ensure scalability. #3 is scalable by nature because the work only describes an information model on a conceptual level. The rest of the related work including the SRA4IACS approach does not cover the topic of scalability. Nevertheless, for

SRA4IACS it would be possible to add measures for scalability, but this is currently not needed due to a low runtime and lightweight implementation.

R-G-5: The works of #4 and SRA4IACS integrate technical evidences, such as specific asset configurations or settings, into the security risk assessment process. The other related work includes only general asset information and no specific asset configurations used as technical evidences.

11.1.2. Information Collection (IC)

Table 11.2 compares and evaluates the five selected most relevant publication series from the related work (see Section 4.3.2) with the expert system from this dissertation for the automation of security risk assessments based on the six defined requirements for the information collection (see Section 6.1). In this category SRA4IACS mainly addresses the deficit D1 of insufficient process coverage (from Section 4.3.3) by defining a complete security risk assessment process and the coupling with safety.

Table 11.2: Evaluation results of the information collection requirements

ID	Requirement	#1	#2	#3	#4	#5	SRA4IACS
R-IC-1	Security Conformity	✔	✔	✔	∼	∼	✔
R-IC-2	Decision Making	∼	∼	∼	∼	✔	✔
R-IC-3	Use Cases	∼	✔	✔	✔	∼	∼
R-IC-4	Safety Coupling	∼	✗	✗	✔	✗	✔
R-IC-5	Resource Consumption	∼	✗	✗	✗	∼	∼
R-IC-6	Data Quality	∼	∼	✗	∼	∼	∼
	Evaluation Score:	**3.5**	**3.0**	**2.5**	**3.5**	**3.0**	**4.5**

R-IC-1: An adequate security conformity is fully provided by the works #1, #2, #3, and SRA4IACS due to an integration of the IEC 62443 standard. The works #4 and #5 only mention a possible usage of the IEC 62443 standard, but do not include it directly.

R-IC-2: All solutions offer at least some aspects to support operator decision making processes. #5 and SRA4IACS comprise an enhanced GUI which can be used to support operators.

R-IC-3: In general, an automated security risk assessment process should be applicable to various use cases to achieve a wide acceptance. The works #2 and #3 cover the most use cases due to their conceptual nature, whereas #4 supports an adequate amount of use cases for the intrusion detection of attackers. #1, #5, and SRA4IACS have only a limited amount of use cases implemented resulting in a smaller area of possible application up to the current status of developments.

R-IC-4: As indicated within Section 3.4, the coupling of security and safety becomes increasingly important similar as the IT/OT convergence. SRA4IACS and #4 integrate safety into the overall methodology as the most important aspect, #1

uses the safety-related ZCR 3.3 as a use case, and #2, #3, and #5 do not integrate safety at all.

R-IC-5: The overall resource consumption of an implementation only plays a minor role. #2 and #3 cannot be evaluated in this regard due to the conceptual level, #4 requires additional efforts due to the learning-based approach, and #1, #5, and SRA4IACS are all lightweight implementations suitable for low-resource hardware typically used within research and development.

R-IC-6: The data quality requirement is not adequately fulfilled by any of the approaches including SRA4IACS due to the manual elicitation of security expert knowledge and missing interoperability among each other. #3 does not contain any defined data at all due to the conceptual level.

11.1.3. Information Formalisation (IF)

Table 11.3 compares and evaluates the five selected most relevant publication series from the related work (see Section 4.3.2) with the expert system from this dissertation for the automation of security risk assessments based on the four defined requirements for the information formalisation (see Section 7.1). The SRA4IACS approach generally addresses the deficit D2 of missing standardised metrics (from Section 4.3.3) by identifying, analysing, and integrating well-accepted and well-established security formalisations for the modelling of information.

Table 11.3: Evaluation results of the information formalisation requirements

ID	Requirement	#1	#2	#3	#4	#5	SRA4IACS
R-IF-1	Credibility	✔	✔	✔	✗	✔	✔
R-IF-2	Knowledge Representation	∼	∼	✗	∼	∼	✔
R-IF-3	Reusability	∼	∼	✔	✔	✗	∼
R-IF-4	Training Data	✔	✔	✔	✗	✔	✔
	Evaluation Score:	**3.0**	**3.0**	**3.0**	**1.5**	**2.5**	**3.5**

R-IF-1: The typically used metrics of complexity and impact for possible security risks are integrated into SRA4IACS and the usage of already defined formalisations enhances the overall result credibility and trustworthiness, similar as in #1, #2, #3, and #5. Only the work #4 uses a training- and learning-based ML approach loosing the overall credibility due to black-box behaviour and missing explainability of results.

R-IF-2: Following the findings of [44], rule-based systems are the most suitable to model security knowledge and experiences. #3 does not define any rules at all, just a general information model in a graphical manner. The SRA4IACS covers the complete security risk assessment process and the associated rules, whereas the other works either only offer a limited scope of knowledge as rules or remain on a conceptual level.

R-IF-3: In general, an information model should be built up in a compositional manner enabling the reusage of certain aspects without affecting other parts. #4 offers an adequate separation of modelling aspects and the reference information model of #3 is divided into four main parts independent from each other. All other related work including SRA4IACS offer a certain degree of reusability by separating information modelling from knowledge processing.

R-IF-4: Generally, the security domain lacks historical and reference data sets to adequately use training- and learning-based approaches. #4 is strictly dependent on available training data to feed the implemented random forest algorithm which need to be acquired in a non-intrusive manner via a network scanner. All other works including SRA4IACS are independent from training data and work on a manually created knowledge base.

11.1.4. Information Usage (IU)

Table 11.4 compares and evaluates the five selected most relevant publication series from the related work (see Section 4.3.2) with the expert system from this dissertation for the automation of security risk assessments based on the three defined requirements for the information usage (see Section 8.1). In this category SRA4IACS mainly addresses the deficit D3 of a low maturity (from Section 4.3.3) by analysing security domain expert knowledge and implementing it as specific and understandable predicate logic rules. This enables a human-readable and machine-interpretable format of knowledge elicitation.

Table 11.4: Evaluation results of the information usage requirements

ID	Requirement	#1	#2	#3	#4	#5	SRA4IACS
R-IU-1	Explainability	✔	∼	✗	✗	∼	✔
R-IU-2	Expert Knowledge	∼	∼	✗	∼	∼	✔
R-IU-3	Rule Characteristics	✔	∼	∼	∼	∼	∼
	Evaluation Score:	**2.5**	**1.5**	**0.5**	**1.0**	**1.5**	**2.5**

R-IU-1: The first requirement is based on the explainability of the achieved security risk assessment results. #3 is a method-agnostic reference model and therefore does not create any results directly. #4 implements a learning-based approach which lacks the possibility to explain specific results due to a black-box like behaviour. #1 and SRA4IACS offer a high degree of explainability based on the manually defined rule set, whereas #5 lacks the presentation of implemented rules.

R-IU-2: According to [44], the modelling of expert knowledge is best done in a rule-based and forward-chained manner. Again, #3 does not define any rules. SRA4IACS offers an adequate amount of rules covering all necessary security risk assessment process steps. The works of #1, #2, #4, and #5 either only implement a narrow scope of knowledge, lack detailed explanation of rule structure, or remain on a conceptual level.

R-IU-3: A similar evaluation can be made for the usage of rules and their characteristics. Only work #1 utilises additional measures for the specification and checking of rules, such as inference or reasoning. All other works including SRA4IACS rely on manual rule specification and checks by a security domain expert.

11.1.5. Information Access (IA)

Table 11.5 compares and evaluates the five selected most relevant publication series from the related work (see Section 4.3.2) with the expert system from this dissertation for the automation of security risk assessments based on the five defined requirements for the information access (see Section 9.1). The SRA4IACS approach generally addresses the deficit D4 of high abstraction levels (from Section 4.3.3) by implementing the AAS concept as a DT into the real-world application scenario of the Customisable Production System from the SmartFactoryOWL in Lemgo.

Table 11.5: Evaluation results of the information access requirements

ID	Requirement	#1	#2	#3	#4	#5	SRA4IACS
R-IA-1	OT Adaptation	✔	✔	✔	∼	✔	✔
R-IA-2	Syntax & Semantics	∼	∼	✔	∼	✔	✔
R-IA-3	Secure Access	✗	✗	✗	✗	✗	∼
R-IA-4	Tool Availability	✔	✗	✗	✔	✔	✔
R-IA-5	Data Sources	∼	∼	∼	∼	∼	✔
	Evaluation Score:	**3.0**	**2.0**	**2.5**	**2.5**	**3.5**	**4.5**

R-IA-1: The access to the specific data about the SUC needs to be adequate for the OT domain. The works of #2 and #3 remain on a conceptual level and therefore can theoretically be integrated into the accessed information model about the SUC. #4 is heavily reliant on network and asset scans which contradicts the desired minimally invasive approach requirement. #1 and #5 use AML files and SRA4IACS uses the AAS for information access. Both formalisms follow a similar approach and the associated discussion is presented in Section 9.2.

R-IA-2: In general, the information shall be represented in a standardised syntax and semantic. SRA4IACS only uses the defined metamodel structure of the AAS which is a de-facto standard. #5 integrates the original style of AML files and #3 completely follows the UML definitions for the information model. The remaining works of #1, #2, and #4 either propose proprietary enhancements to AML or lack the presentation of the implemented syntax and semantic.

R-IA-3: The requirement of secure access to the relevant information is not covered adequately by any of the solutions. The implementations of #1, #4, and #5 do not discuss the topic of additional security mechanisms, whereas SRA4IACS at least provides the possibility of a secure environment based on the utilised software libraries and tools offering secure access functionalities for the AASs.

R-IA-4: The access to the necessary data should be supported by tools. #2 and #3 do not offer any functionality to acquire real-world data due to the lack of implementation. All other works including the SRA4IACS approach offer a tool-based acquisition of information for the security risk assessment based on either AML files or AASs as the industrial DT.

R-IA-5: The last requirement ensures the integration of various heterogeneous data sources into the security risk assessment process. SRA4IACS includes several data sources, such as CVE and CVSS for vulnerabilities, the Intel TAL for attacker characteristics, the MITRE ATT&CK framework for attacker techniques and mitigations, or the IEC 62243 standard for the basics of SLs, FRs, SRs and CRs. The remaining related works only use singular data sources, lack the publication of integrated data sources, or remain on a conceptual level of definitions.

11.1.6. Result Summary

The verification results are shown in Table 11.6 below. A total evaluation score of 23 points is the theoretical maximum for any ideal solution due to the amount of requirements for each category (General: 5 / Information Collection: 6 / Information Formalisation: 4 / Information Usage: 3 / Information Access: 5). The five categories are only used to distinguish between the different content areas represented by the method steps and the amount of requirements per category is not further used for any weighting of the overall evaluation score. Furthermore, all 23 requirements are weighted equally regarding their impact on the overall evaluation score. Neither of the five presented related works nor SRA4IACS reaches the total evaluation score of 23 due to the distinctive advantages and disadvantages. Nevertheless, the SRA4IACS approach is the only one which completely fulfils all nine of the requirements from all categories with a high priority.

Table 11.6: Summary of the verification results based on the requirements analyses

Category	**#1**	**#2**	**#3**	**#4**	**#5**	**SRA4IACS**
General	3.5	1.0	2.0	2.0	2.5	3.5
Information Collection	3.5	3.0	2.5	3.5	3.0	4.5
Information Formalisation	3.0	3.0	3.0	1.5	2.5	3.5
Information Usage	2.5	1.5	0.5	1.0	1.5	2.5
Information Access	3.0	2.0	2.5	2.5	3.5	4.5
Evaluation Score:	**15.5**	**10.5**	**10.5**	**10.5**	**13.0**	**18.5**

The summary is presented here showing how close the related works are among each other with respect to the evaluation score and where each has its strengths and weaknesses. The publication series #2, #3, #4, and #5 all range from an evaluation score of 10.5 to 13.0 representing the solid foundation of related work on which this dissertation is based. The publication #1 achieves an evaluation score of 15.5 due to the close similarity of SRA4IACS with only minor differences in the information

collection and access. The expert system for automated security risk assessments of this dissertation is rated with an evaluation score of 18.5 as the highest one of the comparison representing an important added value.

11.2. Validation: Proposed Solution Approach

The second part of the evaluation covers the validation process of the achieved results regarding the automation of security risk assessments. First, Section 11.2.1 describes the concept to evaluate the new automated expert system from this dissertation by comparing it to a typical manual security risk assessment as a reference set of data. In Section 11.2.2, the two hypotheses are checked and answered regarding the achieved result quality and process automation.

11.2.1. Reference Security Risk Assessment

This subsection aims to describe, compare, and evaluate a typical manual security risk assessment with the automated results from the expert system of this dissertation. Therefore, a reference security risk assessment for the Customisable Production System from the SmartFactoryOWL (as described within Section 5.1.2) is performed manually to allow for a result comparison [196]. The specific SUC for the creation of the reference set of data is situated in the same scope as the defined "Example 2" from Section 10.2 of the prototypical implementation of the expert system focusing on the laser module and the associated modules of conveyor belt, exhaust, and cabinet as a typical application scenario (see Section 1.2). Furthermore, it is the goal to provide a reference security risk assessment also for other related works to have a comparable set of data and information basis which is currently not available at all. This general validation approach can be compared to already available initiatives from other domains, e.g. penetration testing with the Open Worldwide Application Security Project (OWASP) Juice Shop[3], Metasploitable[4], Damn Vulnerable Web Application (DVWA)[5], or buggy web application (bWAPP)[6].

Figure 11.1 compares the general activities to perform security risk assessments either in the manual variant (in white) or the automated variant (in grey). This provides an abstracted overview of the differences and similarities between both variants. The activities generally consist of the two phases of (1) preparation for and (2) performance of security risk assessments. The preparation phase includes gaining experience or building up knowledge as a risk assessor in the manual variant in contrast to defining and modelling the expert system in the automated variant. The performance of the security risk assessment is based on prior art and best practices and executed by a human expert in the manual variant as opposed to the utilisation

[3]https://github.com/juice-shop/juice-shop
[4]https://sourceforge.net/projects/metasploitable
[5]https://github.com/digininja/DVWA
[6]http://www.itsecgames.com

of the prototypical implementation for the result creation and documentation in the automated variant. It is assumed that the preparation phase to provide the necessary knowledge for the security risk assessment is comparable between the manual and the automated variant with regard to required efforts and resources, e.g. time, costs, or personnel [44]. This includes aspects of the manual preparation, such as knowledge and experience gained during studies or daily business, and aspects of the expert system, such as knowledge elicitation and specific information modelling. Due to this estimation, the first phase of preparation is not regarded here for the further evaluation and the focus is set on the performance of security risk assessments to reveal the inherent differences between the manual and automated variants.

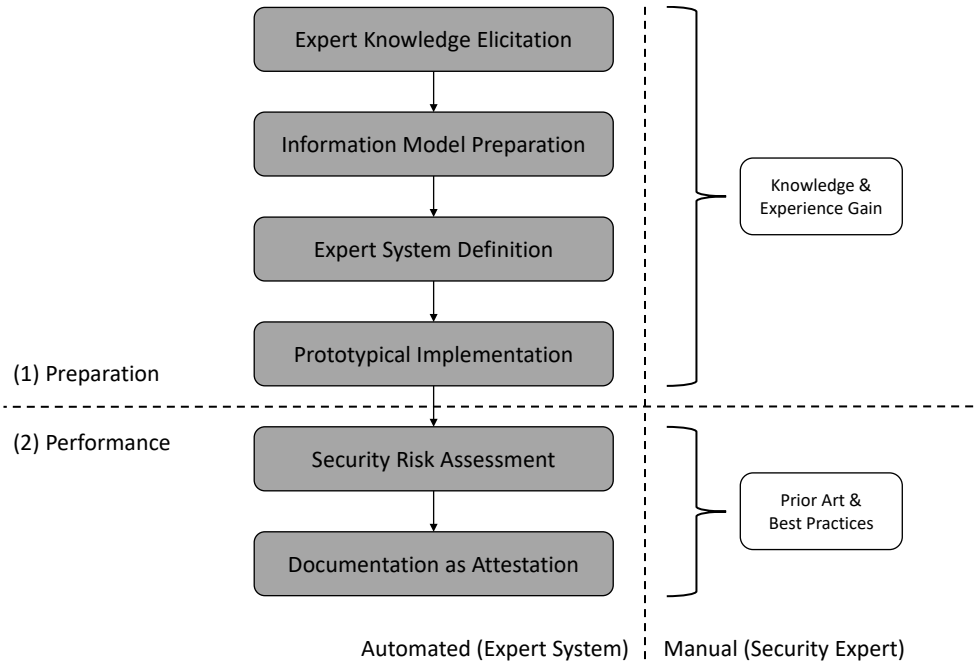

Figure 11.1: Comparison of the general activities to perform manual and automated security risk assessments

The manual reference security risk assessment is prepared by three participants who are further characterised within Table 11.7. The characterisation of the participants is based on their respective job description, role for this evaluation, employer size, background, experience, and the LOK. The presented characteristics are adapted from the work of [44]. Participant A is the author of this dissertation and functions as a typical asset owner of the Customisable Production System as the SUC with the corresponding tasks, such as preparing documentation, granting access to the system, participating in interviews, and finally approving the results. The participants B and C are external OT security consultants and researchers with a sophisticated background and experience. They represent the security risk assessors supporting a typical asset owner with the task of performing the necessary security risk assessment

using their usual methodology without any further instructions, restrictions, or specifications in order to create unbiased results for evaluation.

Table 11.7: Overview and characterisation of the participants for the manual reference security risk assessment

Characteristic	**Participant A**	**Participant B**	**Participant C**
Job Description	Consulting & Research	Consulting	Consulting & Research
Evaluation Role	Asset Owner	Security Risk Assessor	Security Risk Assessor
Employer Size	1-250	1-250	1-250
Background	OT Security	OT & IT Security	OT Security
Experience	1-5 years	$>$ 10 years	$>$ 10 years
LOK	Security Analyst (LOK 3)	Senior Security Analyst (LOK 4)	Senior Security Analyst (LOK 4)

It should be noted that this security risk assessment with only three participants has no statistical relevance but serves the purpose of creating a reference set of data for the comparison. [196]. This is due to the extensive time requirements of a security risk assessment and the overall scarcity of security experts. Nevertheless, the two external security risk assessors (participants B & C) involved are considered to suffice to define the manual reference security risk assessment. As this approach is based on comprehensive expert knowledge and experiences, it can establish trust for the achieved results [196]. The overall findings from the manual reference security risk assessment set of data can be found within Annex C in Section 16.

Table 11.8: Comparison of the manual and automated security risk assessment characteristics (in alignment to the definitions from Section 4.3.1)

Characteristics	**Manual Security Risk Assessment**	**Automated Security Risk Assessment**
Domain	Generally applicable approach	OT-specific approach
Completeness	1-5	1-5
Perspective	Business-driven	Asset-driven
Data Source	1 & 4	1-4
Repeatability	Low capability for repeatability	High capability for repeatability
Timing	Static	Dynamic
Safety Coupling	Independency	Conditional dependency
Network Segmentation	Not performed	Zones & Conduits
Assessment Type	Qualitative	Qualitative
Abstraction Level	Technical and organisational	Only technical
Risk Cardinality	Multiple risks per asset	One risk per asset
Standard Conformity	Not given	IEC 62443
Adoptability	Low	High

In contrast, the automated variant of the security risk assessment is performed by the implemented expert system (as defined within Section 10.2) which is based on the four method steps of this dissertation (see Section 6 to Section 9). Both security risk assessment variants (manual and automated) are based on the same information base consisting of, e.g. asset characteristics, physical connections, network architecture, technical vulnerabilities, and the description of the production process (as defined in Example 2 as the basis for evaluation within Section 10.2.2). Table 11.8 summarises the comparison of the manual and automated security risk assessment characteristics

to understand the differences and similarities between the two underlying procedures. Most of the characteristics are taken from the definitions for the comparison of the related work within Section 4.3.1.

11.2.2. Hypotheses Check

This section covers the checks of the two main hypotheses (H1 & H2) as an important aspect of validation. They are recited below in each subsection and are originally defined within Section 1.3 which introduces and motivates the overall problem statement of this dissertation. The results from the manual reference security risk assessment set of data and the prototypical implementation of the expert system for automation are taken to check and answer the two hypotheses.

Hypothesis 1: Result Quality

> *H1 (Result Quality): If security risk assessments for IACSs are automated, the results are qualitatively comparable to their manual counterpart.*

The first hypothesis (H1) is checked based on distinctive Key Performance Indicators (KPIs). Generally following the definitions from the ISA TR 84.00.09 which are based on the ISO 14253-1 standard, KPIs should be context-specific with a high relevance for the addressed stakeholder, should have a quantitative conformance possibility through expression as a cardinal number or percentage, and their creation and processing should be automated to the maximum extent possible. The specified KPIs used for the further evaluation of this dissertation are summarised within the following list in a defined manner according to the KPI template proposal from the Annex L from the ISA TR 84.00.09. Each KPI has a unique ID for identification and a descriptive text including the content and the naming. Furthermore, the necessary input and the measured output for the KPIs are described. In addition, the source of the KPI is shown providing the capability to relate the KPIs to the associated and relevant documents. Finally, the primary stakeholders as the addressed individuals are specified.

Finding Quantity

- ID: KPI-1
- Description: Quantitative amount of the identified findings based on the security risk assessment process
- Input: Asset inventory
- Output: Total number of findings
- Source: IEC 62443-3-2 → ZCR 2.1 & ZCR 7.1

- Stakeholder: Asset owner

Affected Assets

- ID: KPI-2
- Description: Quantitative amount of the affected assets based on the security risk assessment process
- Input: Asset inventory
- Output: Total number of affected assets
- Source: IEC 62443-3-2 → ZCR 2.1
- Stakeholder: Asset owner

Risk Levels

- ID: KPI-3
- Description: Qualitative levels of the identified security risks for the findings based on the security risk assessment process
- Input: Asset inventory
- Output: Security risk levels for the findings
- Source: IEC 62443-3-2 → ZCR 2.1 & ZCR 7.1
- Stakeholder: Asset owner

The associated results of the automated variant can be found within the presented attestation document from Figure 10.5 in Section 10.2.2 and the results of the manual security risk assessment can be found within Annex C in Section 16. The manual results include six technical (NET-1, VLN-1, VLN-2, VLN-3, VLN-4 & VLN-5) as well as one organisational (ISMS-1) finding. The organisational finding is out of scope for this evaluation because the focus is set on asset-related characteristics as the main source of information for the security risk assessment process to compare the manual and the automated variants. In addition, the technical finding VLN-1 is not taken into account due to the affected MES host asset which is part of the overarching architecture within the public network, but not directly part of the defined SUC in scope. This finding was only revealed during the manual variant because of an extended and optional coverage. Table 11.9 summarises the comparison of the three defined KPIs regarding the manual and automated security risk assessments.

KPI-1 describes the amount of identified findings for the SUC in total. The manual security risk assessment process identified five findings, whereas the automated

Table 11.9: Comparison and evaluation of the KPIs regarding the manual and automated results of the security risk assessments

ID	Name	Manual Result	Automated Result
KPI-1	Finding Quantity	5	2
KPI-2	Affected Assets	7	2
KPI-3	Risk Levels	Medium	Very High

variant revealed two findings. KPI-2 covers the amount of affected assets of the SUC documented within the findings during the security risk assessment itself. The expert system automatically identified two affected assets and the manual variant results in seven affected assets. KPI-3 describes the evaluated levels of security risks for the SUC which are then further used for the subsequent decision making processes by, e.g. the operator or the asset owner. Here, the manual process revealed only *"Medium"* security risks, whereas the automated process identified only *"Very High"* security risks. Figure 11.2 summarises and compares the results based on the three KPIs regarding finding quantity, affected assets, and security risk levels. The provided heatmap is based on the presented SUC architecture from Section 10.2.2. The blue circles indicate the affected assets from the automated security risk assessment, the black circles document the affected assets from the manual process. In addition, the amount of findings can be seen by the red boxes attached to the affected assets which also include the associated risk levels.

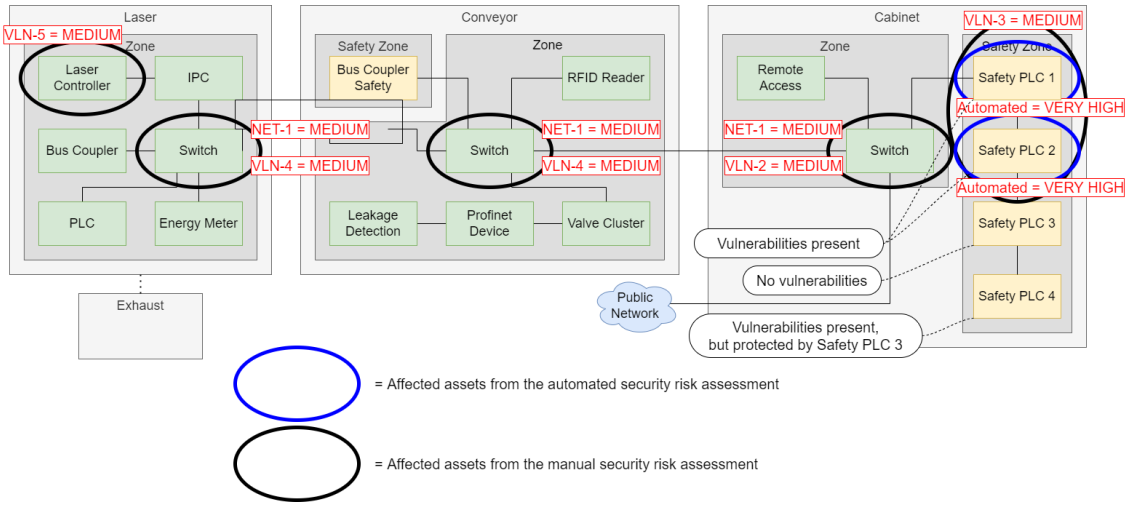

Figure 11.2: Overview of the result comparison based on the three KPIs regarding finding quantity, affected assets, and security risk levels

The results within Table 11.9 and Figure 11.2 reveal certain differences between the manually created reference set of data and the automated security risk assessment process regarding the three KPIs which are discussed in the following paragraphs.

KPI-1 covers the finding quantity. The automated expert system identifies target assets for the security risk assessment process based on the network segmentation beforehand and the inherited safety characteristic related to a direct usage of the asset for the production process. This currently limits the scope of assets by the automated variant to purely typical OT assets, such as PLCs or bus couplers which are utilised within safety functions. In contrast, typical IT assets, such as networking devices, e.g. switches or routers, or desktop computers, are out of scope for the automated security risk assessment due to the underlying information model and security risk determination. Nevertheless, these assets can also be safety-relevant based on their usage as the manual reference set of data shows, e.g. by communication via safety protocols as it is the case for the three switches or as a host for control software with safety implications as it is the case for the laser controller. This explains the different amount of findings for the two variants. In addition, the manual finding VLN-3 also includes the safety PLC #4 which is documented as not affected ("No Relevant Risk") by the automated variant. This difference originates from the varying inspection of the bridged safety PLCs. The automated risk assessment assumes the communication protocol to be secure, e.g. regarding encryption or authentication, and therefore does not list PLC #4 as an affected asset, because it is protected by the PLC #3 prior within the the line topology. An extension of the automated security risk assessment process towards an increased scope of assets is already in planning and part of the future work. This will address and resolve the current discrepancy in the finding quantity. However, the overall achieved result quality of the automated expert system is still adequate with regard to the findings, only the scope is more narrow at the moment.

KPI-2 describes the amount of affected assets of the SUC. The automated security risk assessment process always creates one finding per identified asset resulting in a ratio of 1:1. Whereas the manual process has a flexible ratio of assets to findings which results in the varying numbers and even a higher number of affected assets (7) in comparison to the total amount of findings (5). This leads to the different amount of affected assets of the same SUC in comparison between the manual and the automated variant. In addition, this distinction is also related to the described scope within KPI-1 beforehand and will be addressed with the mentioned upcoming activities resulting in no further implication for the result quality.

KPI-3 addresses the security risk levels of the findings per affected asset. The automated variant resulted in only *"Very High"* risks, whereas the manual reference set of data documents only *"Medium"* risks. This discrepancy is due to the usage of different risk matrices for the final risk determination based on the underlying complexity and impact metrics. Figure 11.3 shows the two different risk matrices. As also presented within Annex B of the IEC 62443-3-2 standard, various risk matrices are widely available and are typically utilised for the security risk assessment processes dependent on the associated use case. In this case, the risk matrix used for the manual variant is based on experiences, the prior art, and known best practices of the security expert to match technical as well as organisational findings. This results in the shown 4x3 matrix on the left side. In contrast, the risk matrix used for the

Risk matrix used for the manual security risk assessment

Complexity	Impact		
	Low	Medium	High
Very High	Low	Low	Low
High	Low	Low	Medium
Medium	Low	Medium	High
Low	Medium	High	High

Risk matrix used for the automated security risk assessment

Complexity	Impact		
	None (N)	Low (L)	High (H)
High (H)	Very Low	Medium	High
Low (L)	Low	Medium	Very High

Figure 11.3: Comparison of the different risk matrices for the manual and the automated security risk assessment process

automated expert system is solely intended for technical findings and is therefore based on the CVSS characteristics of vulnerabilities regarding complexity and impact resulting in the 2x3 matrix on the right side. This difference is generally typical for security risk assessments performed by various stakeholders and can also be found within real-world examples, e.g. performances done by different security experts, teams, or companies. Consequently, the evaluation regarding the result quality of the automated security risk assessment is not influenced.

> *Under the remarks formulated for KPI-1, KPI-2, and KPI-3, the hypothesis H1, which covers the result quality, is valid.*

Hypothesis 2: Process Automation

> *H2 (Process Automation): If the automation degree of security risk assessments for IACSs is increased, the overall required efforts for the operator are reduced.*

Table 11.10 shows the comparison of the manual and automated security risk assessment processes mapped towards the ZCRs from the IEC 62443-3-2 standard (as described within Section 3.6). This comparison documents the specific security risk assessment tasks which are performed for the manual and the automated variant. The comparison is done via a three-tier scale: ✔= ZCR actively performed / $\sim$ = ZCR passively addressed / ✗= ZCR not done at all. The manual reference set of data as a typical security risk assessment covers nine ZCRs in total and required 19.0 hours of efforts consisting of the tasks by a typical asset owner (10.0 hours) and a typical risk assessor (9.0 hours).

The automated expert system actively covers six ZCRs (indicated by the ✔symbol) as originally specified within Section 5.2 which make up nearly 50% of the required efforts as revealed at the beginning of this dissertation. Furthermore, nine additional

Table 11.10: Comparison of the ZCRs from the IEC 62443-3-2 standard towards the manual and the automated security risk assessment steps

RM Steps	**Manual Coverage**	**Asset Owner [h] (Manual)**	**Risk Assessor [h] (Manual)**	**Automated Coverage**
ZCR 1.1	✔	3.0	1.0	∼
ZCR 2.1	✔	4.0	2.0	✔
ZCR 3.1	✗	0.0	0.0	✔
ZCR 3.2	✗	0.0	0.0	✗
ZCR 3.3	✗	0.0	0.0	✗
ZCR 3.4	✗	0.0	0.0	✗
ZCR 3.5	✗	0.0	0.0	✗
ZCR 3.6	✗	0.0	0.0	✗
ZCR 4.1	✗	0.0	0.0	✗
ZCR 5.1	✔	1.0	0.5	✔
ZCR 5.2	✔	1.0	1.0	✔
ZCR 5.3	✔	0.0	0.5	∼
ZCR 5.4	✔	0.0	0.5	∼
ZCR 5.5	✔	0.0	0.5	∼
ZCR 5.6	✗	0.0	0.0	✔
ZCR 5.7	✗	0.0	0.0	✗
ZCR 5.8	✗	0.0	0.0	✗
ZCR 5.9	✗	0.0	0.0	✗
ZCR 5.10	✗	0.0	0.0	✗
ZCR 5.11	✗	0.0	0.0	✗
ZCR 5.12	✗	0.0	0.0	✗
ZCR 5.13	✔	0.0	2.0	∼
ZCR 6.1	✗	0.0	0.0	✗
ZCR 6.2	✗	0.0	0.0	∼
ZCR 6.3	✗	0.0	0.0	✗
ZCR 6.4	✗	0.0	0.0	∼
ZCR 6.5	✗	0.0	0.0	✗
ZCR 6.6	✗	0.0	0.0	∼
ZCR 6.7	✗	0.0	0.0	✗
ZCR 6.8	✗	0.0	0.0	✗
ZCR 6.9	✗	0.0	0.0	∼
ZCR 7.1	✔	1.0	1.0	✔
Total:	**9 ZCRs**	**10.0**	**9.0**	**6 / (15) ZCRs**

ZCRs are passively addressed (indicated by the ∼ symbol), e.g. the SUC description for the ZCR 1.1., the usage of complexity and impact to calculate risks for the SL-T analysis covering ZCRs 5.3 to 5.5, or the fulfilled documentation requirements due to the internal expert system information processing including ZCRs 5.13, 6.2, 6.4, 6.6, and 6.9 in addition to the original scope. This further reduces efforts by covering

additional ZCRs in a faster manner with a comparable and adequate result quality. Following this analysis, the automated security risk assessment process covers all the ZCRs from the manual reference either actively or passively and six additional ones. This represents a sufficient and practically-usable security risk assessment process coverage for the automated expert system as one main result of this dissertation.

In addition to the comparison of ZCRs, Table 11.11 shows the evaluation of the manual and automated security risk assessment on an abstracted level of details. The developed expert system increases the available LOA to 4 and reduces the needed LOK to 1 resulting in an overall improved usability of the process. The efforts to perform a security risk assessment for the first time are high and comparable between the manual and the automated approach (in alignment with the estimations from [44] and Figure 11.1 before). The automated expert system can achieve an additional improvement with regard to the repetition of security risk assessments for the same SUC reducing the needed efforts to a low level in comparison to a medium level for the manual process.

Table 11.11: Evaluation of the needed efforts for manual and automated security risk assessments (in alignment with the estimations from [44])

Performance	**Manual**	**Automated**	**Evaluation Description**
First-time	High	High	Comparable, but different type of efforts
Repeated	Medium	Low	Improvement plus keeping result quality
Available LOA	Max. 3	4	Increase of tool-based support
Needed LOK	Min. 3	1	Reduction of required skill level

The runtime of the expert system algorithm is below 10s on every computing environment tested so far ranging from embedded devices and tablets to business laptops. In comparison, the creation of the manual security risk assessment reference set of data took 19 hours (see Table 11.10). This reveals that the usage of the automated expert system results in significant time savings with regard to the performance of security risk assessments. Nevertheless, it should be noted that this result only applies to the investigated Customisable Production System demonstrator as the SUC so far and was not repeated on any other SUC yet. Currently, it is estimated that the algorithm has a linear runtime development within the range of $<$20 assets and $<$20 network connections as proven by the investigated SUC. Any statements regarding the algorithm runtime for further SUC sizes and complexities cannot be made yet due to missing demonstrators and the associated modelling of information, such as the specification of the corresponding AASs. Furthermore, a human security risk assessor would never perform a security risk assessment of a too large and complex SUC completely manually, but instead will resort to more high-level analyses and aggregation mechanisms. This will result in a valid approximation of the SUC in an adequate time for the manual variant. A comparable functionality is currently missing for the expert system and will be part of the future work. However, additional achievements of the automated expert system include

the results of the security risk assessment process being completely traceable and remaining on a consistent level of quality due to the independence of the security risk assessors' personal knowledge, experience, and procedures. These characteristics further emphasise the adequate automation for the security risk assessment process by the automated variant.

> *Under the remarks formulated for Table 11.11 and Table 11.10, the hypothesis H2, which covers the process automation, is valid.*

11.3. Credibility: Outcome Interpretation

The third part of the evaluation covers the aspect of outcome interpretation. The verification and validation beforehand ensure the basis for a high credibility of the achieved results. Subsequently, this section summarises the learnings of this dissertation based on the discussion of the generalisability of the automation of security risk assessments (see Section 11.3.1) and regarding the concluded recommendations for a practical relevance (see Section 11.3.2).

11.3.1. Generalisability

The discussion and clarification regarding the generalisability of the results are important steps for the overall evaluation. Therefore, this section contains three parts to discuss the topic of generalisability based on the following three guiding questions:

1. Reusability: What are the specific aspects utilised within the four main method steps and are they generally reusable within other works as well?
2. Suitability: Does the chosen Customisable Production System from the SmartFactoryOWL represent a typical IACS and does it fit as a SUC?
3. Transferability: Is it possible to transfer the resulting expert system towards different systems, problem statements, domains, and lifecycle phases?

Guiding question #1 (Reusability): Table 11.12 shows the evaluation of the four conceptual method steps of information collection, formalisation, usage, and access (see Section 6 to 9) and the specific prototypical implementation of this dissertation (see Section 10) with regard to the generalisability. The already utilised three-tier scale (✔= aspect is fully reusable / ∼ = aspect is partly reusable / ✗= aspect is not reusable) is used to qualitatively assess the reusability of the general ("G") and specific ("S") aspects of each method step. This can then be translated into a first estimation of the degree of generalisability of the overall results. Principally, general concepts always need to be specified to be able to fulfil a certain

purpose [79], in this case to create an expert system for the automation of security risk assessments. The method step of information collection is conceptually based on the definition of the swimlanes representing the security risk assessment steps in scope and characterising them with further details. The approach is highly reusable due to its technology- and implementation-neutral manner. In contrast, the specific scope of the six focused ZCRs from the IEC 62443-3-2 standard limits the reusability from a content point of view to a certain degree. By using UML class diagrams for the information formalisation method step to create a semi-formal model, a high reusability can be achieved. The specific identification, analysis, and integration of the chosen security formalisms, such as the MITRE ATT&CK framework or the Intel TAL, creates the specific implementation of this dissertation. This slightly reduces the overall reusability due to the chosen formalisms possibly not meeting the expectations, preferences, or requirements from other stakeholders. Nevertheless, all of the utilised security formalisms are non-proprietary, publicly available, widely spread, and well-acknowledged within the security domain. The method step of information usage is based on predicate logic to elicitate the human expert knowledge into rules. On the one hand, the general logic-based concept is completely reusable for other works, and on the other hand, the specific set of defined rules is also reusable due to the focus on the IEC 62243-3-2 ZCRs as a generally accepted process for security risk assessments. The last method step of information access utilises the concepts from the AAS by proposing a submodel template as a basis for the information model required for this dissertation. These submodel templates represent a general blueprint for other works and can be widely reused. Nevertheless, the specifically created AASs solely represent the Customisable Production System from the SmartFactoryOWL in Lemgo as a SUC and cannot be reused directly towards other systems without adaptations.

Table 11.12: Evaluation of the four main method steps regarding their general aspects of the conceptual level and their specific aspects of the prototypical implementation

Method Steps	**General Aspects**	**Specific Aspects**	**Reusability**
Information Collection	Swimlanes (BPMN)	Scope of six main ZCRs	G: ✔/ S: ∼
Information Formalisation	UML class diagrams	Integration of security formalisms	G: ✔/ S: ∼
Information Usage	Predicate logic	Rules for security risk assessments	G: ✔/ S: ✔
Information Access	AAS submodels	AASs representing the SUC	G: ∼ / S: ✗

Guiding question #2 (Suitability): Following the definitions from Section 1.4, the Customisable Production System from the SmartFactoryOWL can be seen as a typical IACS as defined within the IEC 62443-1-1 standard. Furthermore, it fulfils the RMS characteristics from Section 1.2 enabling the desired application scenario. It includes a collection of personnel, hardware, software, procedures, and policies enabling safe, secure, and reliable operational and maintenance capabilities. The Customisable Production System can be characterised by a medium level of complexity and size, utilises a variety of components from different manufacturers, and has a

real-world impact of safety-related issues regarding the integrated laser engraving machine (see Sections 10.2 and 5.1). In addition, it represents an automation solution including essential functions (as defined within the IEC 62443 standard[7]), such as for control, safety, or complementary tasks. These characteristics lead to the utilisation of the Customisable Production System as an adequate SUC for the automation of security risk assessments.

The integrated and implemented assets of the Customisable Production System also belong to the typically utilised assets within the OT domain, such as network devices (27.8%), PLCs (14.9%), or HMIs (7.0%) [10], and are produced by common component manufacturers, e.g. Phoenix Contact or Siemens [230]. There are also practical reasons originating from this dissertation favouring the Customisable Production System: Availability of information (already partly in the form of AASs), established contacts to the responsible stakeholders, and a general accessibility of the system and the facility. All the mentioned characteristics of the Customisable Production System as the SUC of the specific implementation of this dissertation ensure a high degree of generalisability regarding the achieved results.

Guiding question #3 (Transferability): The specified and implemented expert system as a result of this dissertation achieves a TRL of 6, a LOA of 4, and requires a LOK of 1 (see Section 10.2). Furthermore, it combines functionalities from different tool categories defined within Section 4.1: Checklist, OT monitoring tool, and documentation. In general, this increase in the development degree, the higher automation, and the reduction of the needed knowledge, enable a high adoptability and repeatability as defined within Section 4.3.1 for the characterisation of security risk assessment processes. By doing so, access to and usage of the expert system is simplified resulting in a possibly wide usage by various industrial stakeholders [137]. In addition, the separation of the expert system into information modelling and knowledge processing provides a transferability of certain parts into other works (as described within Section 4.2.2). This also holds true for all automated activities from Figure 11.1 to perform security risk assessments.

The Customisable Production System from the SmartFactoryOWL as a SUC represents an IACS with medium complexity due to the present amount of assets, the variety of component manufacturers and vendors, the network topology, and the amount of modules. Nevertheless, a scalability of the overall solution to more complex IACS is given due to the transferable general aspects of each method step. This includes the ability to handle complex IACS by an adequate level of abstraction to achieve practical and usable results [37]. To achieve that, the information modelling would need an extension based on prepared AASs, but the knowledge processing based on the rules could be reused directly. This creates additional one-time efforts for the information modelling which is dependent on the SUC size and complexity, but afterwards the new and updated expert system can be used endlessly for different systems with reduced efforts. This results in a gradually improving effort to performance ratio depending on the amount of usages of the

[7] https://isagca.org/isa-iec-62443-standards

expert system during the long term. Furthermore, the presence of already specified DTs as information sources are generally increasing at the moment and consequently the efforts to model a certain SUC decreases as well [213].

The presented expert system for automated security risk assessments with a focus on safety could also be utilised for other problem statements. The general aspects of the solution (see Table 11.12) can be adapted and reused for the information modelling and information processing regarding additional ZCRs, similar as within the work of [146]. This would only require a manual translation of the IEC 62243-3-2 standard and expert knowledge elicitation into the predicate logic rules based on a clarification of the necessary data for the rules within the swimlanes, the UML class diagrams, and finally within the AASs. Other parts, e.g. the documentation via the attestation and the general software architecture, can be kept unchanged due to the corresponding similarities towards [146] as already discussed within the related work in Section 4.3.2.

A transfer of the overall results of this dissertation is theoretically possible towards other domains with similar requirements of security risk assessment processes, such as critical infrastructure, building automation, transportation, or smart grids, but needs to be analysed and evaluated within the future work in a more detailed manner [44]. The currently chosen and focused security objective of safety results in a safety-informed security approach (as defined within Section 3.4). Nevertheless, the security objective can be exchanged, e.g. to availability, by adapting the associated asset characteristics which are used to determine the targets of the security risk assessment within the SUC. This would require changes towards the information modelling and the knowledge processing aspects of the expert system from this dissertation.

Finally, the transferability towards additional lifecycle phases could be achieved. This is solely dependant on the presence of the corresponding AASs for the specific lifecycle phase. It would be possible to support, e.g. the design and engineering of IACSs with the proposed expert system. This would enable well-informed decisions regarding further adaptations of an implementation or additional mitigation measures based on the results of the automated security risk assessment already during the development phase. The foreseen security-enhancing DTs from [213] would be able to support these kind of activities as well.

11.3.2. Concluding Recommendations

The last part of the overall evaluation covers the identified deficits. Furthermore, the main learnings are presented at the end of this section. First, Table 11.13 summarises the overall evaluation results regarding the achieved degree of improvement, the corresponding method steps for information and process modelling (Section 6 to 9), and the proposed solution concepts to address deficits of this dissertation (from Section 4.3.3).

The first deficit D1 summarises the insufficient process coverage regarding a complete and usable security risk assessment. The method step of information collection (see Section 6) addresses this issue by scoping the focus of this dissertation

Table 11.13: Addressing of the identified deficits by the solution approach of this dissertation based on the associated method steps and concepts

Identified Deficit	Improvement	Method Step	Proposed Solution Concept
D1: Insufficient process coverage	Reduced (∼)	Information Collection	Swimlanes focusing on the six most resource-intensive ZCRs
D2: No standardised metrics	Fixed (✔)	Information Formalisation	Alignment and compliance to the concepts from the IEC 62443 standard
D3: Low approach maturity	Reduced (∼)	Information Usage	Prototypical implementation containing elicited expert knowledge
D4: High abstraction level	Fixed (✔)	Information Access	Industry-grade demonstrator as the SUC based on specific AASs

around the six main ZCRs identified as the most resource-intensive ones. In addition, the defined process within the swimlanes covers all important aspects to achieve an initial security risk assessment enhanced with additional aspects, e.g. threats or vulnerabilities (see Section 5.2). Nevertheless, the resulting scope of this dissertation is limited to the chosen ZCRs in focus. This leads to a partial reduction as the improvement of the deficit D1. The deficit D2 includes the finding that the integration of standardised metrics is missing. The method step of information formalisation (see Section 7) fixes this by achieving alignment and compliance to the IEC 62443 standard and the associated definitions of concepts, e.g. the various types of SLs or the FRs, SRs, and CRs. The developed expert system based on the prototypical implementation contains the elicited knowledge from the human security domain experts translated into the logic-based rules (see Section 8). This enables the reduction of the deficit D3 regarding the common low maturity based on the achieved TRL of 6. The last deficit D4 describes the generally high abstraction level of the related work mainly remaining on conceptual levels and lacking credible evaluations. The fourth method step of information access covers this deficit by specifying AASs representing the Customisable Production System from the SmartFactoryOWL as an SUC (see Section 9).

The second part of this section covers the main learnings gained during the course of this dissertation. These learnings follow a certain descriptive structure consisting of a theoretical cause-effect relationship ("assumption"), the gained knowledge during the course of this dissertation ("conclusion"), the confirmation degree ("assurance"), and the practical design and action recommendations ("guidance"). The following list summarises the main learnings and provides the final insights into the outcome interpretation of the evaluation regarding credibility:

#1: Security Compliance

- Assumption: Security risk assessment processes can be defined in a practical manner compliant to the IEC 62443-3-2 standard [37].
- Conclusion: The swimlane definition and the prototypical implementation are based on the three types of SLs (Capable / Target / Achieved), the five SL values (0-4), the seven FRs, and the list of SRs and CRs from the IEC 62443-3-2 standard.

- Assurance: High → Section 6.2
- Guidance: The current development status of the IEC 62443-3-2 standard allows for a practical integration of the associated concepts into security risk assessment processes and should be used by SMEs.

#2: Safety Coupling

- Assumption: The alignment of security and safety into one goal-oriented approach can be achieved [37].
- Conclusion: Analysis of assets which are suitable for safety via their specific characteristics and integration of safety as the main security objective of the security risk assessment.
- Assurance: High → Section 3.4
- Guidance: The coupling of security and safety is generally possible and is achieved in the form of a safety-informed security approach. A combined approach for holistic co-engineering is not recommended at the moment due to the loss of focus and details from both domains.

#3: Automation Degree

- Assumption: Parts of the overall security risk assessment can be automated based on a traceable and repeatable process [37].
- Conclusion: The collected and formalised information plus the rule-based knowledge elicitation are used within a semi-formal model as the basis for the prototypical implementation.
- Assurance: High → Section 4.2
- Guidance: The additional manual efforts regarding the identification, analysis, and specification of the automated security risk assessment are amortised by the improved usage to generate traceable results in a repetitive manner afterwards. This results in a gradually improving effort to performance ratio depending on the amount of usages of the expert system during the long term.

#4: Knowledge Elicitation

- Assumption: Security expert knowledge from the OT security domain can be collected, formalised, and represented best in a rule-based format [44].
- Conclusion: The analysis of the related work and the state of the art revealed a wide usage of rule-based knowledge. Furthermore, the utilisation of predicate logic fulfils all defined requirements for the expert system developed within this dissertation and allows the separation of information modelling and information processing.

- Assurance: Medium → Section 8.2
- Guidance: Security expert knowledge should be elicited into a rule-based format to make it human-understandable and machine-readable. The additional usage of inference and reasoning functionalities needs to be further investigated.

#5: Usability

- Assumption: Automated security risk assessments can be practically usable via an adequate result abstraction for the user in a graphical manner.
- Conclusion: The sophisticated approach to determine the security risk assessment results is summarised within the attestation PDF document containing all relevant information and hiding details which are not important for the subsequent decision-making process. Nevertheless, an interactive and graphical GUI is currently missing. This needs additional focus in the future. Furthermore, all contents regarding the prototypical implementation are available in a public GitHub repository to be tested and used.
- Assurance: Low → Section 10.2
- Guidance: Expert systems definitely require an adequate GUI for the respective users and their experience. In addition, the utilised software needs to be traceable and credible. Therefore, an open source publication of the code is highly advised.

Part V.

Epilogue

12. Summary

Security risk assessments are generally performed too rarely or even not at all, especially within SMEs and regarding manufacturing systems in the OT domain. This is mainly due to the extensive cost and time requirements for the associated tasks, which currently need to be conducted manually by security experts. Therefore, this dissertation proposes a method for information and process modelling towards the automation of security risk assessments. The underlying hypotheses suggest that this method will achieve qualitatively comparable results and will reduce the required manual efforts for the security experts.

The analysis of the fundamental background and the state of the art highlights several important developments, such as the IT/OT convergence and the coupling of security & safety, as well as a dynamically changing security landscape, based on technical vulnerabilities, rising threats, occurring incidents, and sophisticated human adversaries. To tackle these challenges, a plethora of commercial tools and approaches from the research domain are available to perform security risk assessments for IACSs. However, four main deficits regarding the automation of security risk assessments are identified within the state of the art: Insufficient process coverage, no standardised metrics, low maturity, and high abstraction levels. In addition, the practical and theoretical analysis of three industrial-grade demonstrators from the SmartFactoryOWL in Lemgo exposed six ZCRs (2.1, 3.1, 5.1, 5.2, 5.6, and 7.1) as the most resource-intensive tasks from the security risk assessment process standardised within the IEC 62443-3-2. These tasks are focused within this dissertation for the increase of the LOA and the reduction of required LOK.

The aim of the method for information and process modelling towards the automation of security risk assessments is the definition of an expert system. This is accomplished through the four main method steps of information collection, formalisation, usage, and access. The results include a conceptual framework and a prototypical implementation for the automation of security risk assessments designed for the use by system integrators and asset owners within the OT domain. The developed expert system covers the complete process of assessing security risks: Preparation and scoping, risk identification, analysis, and evaluation as well as documentation. In addition, the coupling towards the safety domain is ensured by adopting an asset-driven and integrating safety-informed security approach. This underscores the feasibility of coupling security and safety measures, while highlighting the current impracticality of holistic co-engineering due to a potential loss of focus and detail in both domains. The associated prototypical implementation SRA4IACS reaches a TRL of 6 and is published in an open source manner within a GitHub repository to enable a high degree of adoptability and reusability.

The first method step of information collection uses swimlanes, as a lightweight version of the BPMN, to specify the underlying security risk assessment process as the basis for automation. This can be interpreted as an implementation of the theoretical process outlined in IEC 62443-3-2 addressing the first deficit of insufficient process coverage. In addition, it clearly demonstrates that the current development status of the IEC 62443-3-2 standard allows a practical integration of the associated concepts into security risk assessment processes and should be used by SMEs for their manufacturing systems.

Afterwards, the second method step of information formalisation covers the integration of already established and acknowledged frameworks into the security risk assessment process in a technology- and implementation-agnostic way. By doing so, an object-oriented and interoperable semi-formal information model is defined as the basis to automate security risk assessments and the second deficit of missing standardised metrics is focused on. This approach slightly limits the overall generalisability due to the additional manual efforts required for the identification, analysis, and specification of the automated security risk assessment. Nevertheless, they are amortised by the improved efficiency in generating traceable results consistently afterwards. In addition, the more frequently the expert system is utilised later on, the better its performance ratio between manual efforts for the initial specification and the overall achieved improvements becomes over time.

Furthermore, the third method step of information usage includes the elicitation of security expert knowledge into rules based on predicate logic to address the third deficit of low maturity. This enables a high degree of reusability and facilitates the automated security risk assessment to accomplish credibility and trustworthiness. The results of this research evidently prove that security expert knowledge should be elicited into a rule-based format to make it human-understandable and machine-readable. The additional usage of inference and reasoning functionalities for the logic rules will be further investigated.

Finally, the fourth method step of information access describes the translation of the developed information model into the AAS data structure and proposes a submodel that enables the information usage for the automation of security risk assessments. This implementation increases the overall usability of the approach and concentrates on the fourth deficit of high abstraction levels. Nevertheless, it can be further expanded as expert systems should include a GUI designed to support relevant users with their dedicated level of experience and their subsequent decision making processes. In addition, the utilised software needs to be traceable and credible. Therefore, an open source publication of the source code is highly advised.

The overall evaluation of the achieved results towards the automation of security risk assessments is based on the three aspects of verification, validation, and credibility. An extensive analysis of requirements is used to verify the developed expert system. The 23 identified requirements are evaluated for the SRA4IACS approach and compared to the five most relevant publication series in the related work. This shows that the SRA4IACS expert system fulfils all of the highly prioritised requirements and achieves the highest overall evaluation score. The validation is based on the

creation of a manual reference security risk assessment by two security experts and the comparison with the results from the automated security risk assessment by the expert system. This comparison shows a lower need regarding manual efforts and requirements for the automated security risk assessment due to the increased LOA to 4 (from max. 3 before) and the reduced LOK to 1 (from min. 3 before). In addition, it is shown that the result quality of the expert system is comparable to the manual security risk assessment and even surpasses the capability to create credible and consistent results. The validation finally proved the two main hypotheses of this dissertation regarding result quality and process automation as valid. In the end, the evaluation documents the credibility of the achieved results based on the three aspects of reusability, suitability, and transferability.

In conclusion, the contribution to the automation of security risk assessments for the OT domain extends the state of the art in research and technology. In addition, the achieved results revealed further challenges, which are relevant for this field of research and are shown within the next section.

13. Future Work

Based on the summarised results of this dissertation presented before, several questions still remain open and further improvements are possible. These are structured within three main research directions: (1) Safety coupling, (2) technological advances, and (3) knowledge transfer. Each of the research directions is outlined below as an outlook towards the possible future work.

(1) Safety coupling: The further alignment of the automated security risk assessment towards safety still presents significant potential for improvements. Currently, the theoretical suitability of assets for safety functions is integrated as an asset characteristic, which is then used for network segmentation and for identification of target assets for the security risk assessment. This state could be enhanced by two approaches. First, responsible stakeholders, such as the system integrator or the asset owner, could maintain the specific information model regarding the real practical usage of assets for safety functions in contrast to their theoretical suitability. Second, other assets within the network architecture may also have an impact on the overall safety of the system, e.g. switches, routers, or industrial computers. These safety-relevant assets should be integrated into the security risk assessment as well.

Another possible addition aims at the utilisation of the results from the documented security attestation. The effect on safety functions is still under discussion, because the result of the automated security risk assessment could impact the system in near real-time, resulting in continued normal operation, degraded functionality, full system stoppage, or emergency shutdown.

The last aspect for improvement regarding the safety coupling involves the prioritisation of the used attack techniques from the MITRE ATT&CK ICS framework. Currently, all possible attacker techniques are integrated with the same priority. However, initial approaches are already available within the research domain[1] based on attacker campaign analyses, by VERIS[2] for confirmed real-world incidents, by MITRE[3] regarding available kill chain information, or within the work of [185] based on intrusion detection. These should be checked and integrated if feasible.

(2) Technological advances: At present, the automated expert system for security risk assessments focuses on Ethernet-based communication paths and line topologies within OT networks. Thereby, serial, proprietary, or bus-based communication protocols and other network architectures, such as ring, have not been adequately addressed thus far. This could be resolved by the analysis and integration

[1] https://cyote.inl.gov
[2] https://verisframework.org
[3] https://attack.mitre.org/resources/sightings

of the additional communication capabilities and the associated implications for the defined information model and prototypical implementation. Nevertheless, the basis for this extension is already given and currently in discussion.

Furthermore, additional security frameworks and formalisms could be integrated to enhance the current state, which are especially further developed by MITRE as one of the main key hubs for the security community worldwide: MITRE D3FEND[4] as a knowledge graph for countermeasure selection, MITRE campaigns[5] to track intrusion activities, MITRE EMB3D[6] as a threat model for embedded devices within critical infrastructure, or Caldera Pathfinder[7] as an automated vulnerability scanner.

In addition, the communication capabilities and interfaces for the information access of the SUC could be extended. First, the current usage of a passive AASs should be further developed based on the available concepts for re-active or pro-active AASs to reflect the future needs for communication. Second, other approaches currently in development, such as the CSAF to exchange security advisories or the SCAP for automated compliance evaluation, could be used to enhance the amount of available and integrated information for the security risk assessment and would improve the overall generalisability and applicability of the results due to standardised interfaces.

The last aspect of technological advances addresses the domain of AI via ML. This is currently not regarded due to missing training data and the explicit need for complete transparency and explainability of the results. Nevertheless, this is already integrated into further research activities within the succeeding SUSI project[8] by testing the adequate usage of large language models, few shot learning, and online learning.

(3) Knowledge transfer: The security domain itself, as well as the IEC 62443 series of standards, is subject to constant changes and further developments. The currently utilised concepts and metrics from the IEC 62443 represent the basis of the available contents.

Firstly, in addition to the technical metrics currently used, other metrics, such as the maturity level to describe procedural security and the protection level summarising technical and organisational security aspects, could be integrated into the expert system for the automated security risk assessments.

Secondly, further evaluation criteria regarding CRs are currently defined within the upcoming IEC 62443-6-1 and regarding service provider requirements within the IEC 62443-6-2. The developed and implemented status at the moment should be investigated with the new contents from upcoming versions of the standard once publicly available.

[4] https://d3fend.mitre.org
[5] https://attack.mitre.org/campaigns
[6] https://www.mitre.org/news-insights/news-release/mitre-red-balloon-security-and-narf-announce-emb3d
[7] https://github.com/center-for-threat-informed-defense/caldera_pathfinder
[8] https://www.init-owl.de/en/research/projects/detail/software-basierte-unterstuetzung-von-security-risikobeurteilungen-in-der-industrie

Thirdly, future versions of the IEC 62443[9] will include profiles to adopt a defined set of requirements for a certain use case. The results of this dissertation could be the basis for a profile regarding the automated security risk assessment for modular IACSs.

The fourth proposal for improvement of knowledge transfer addresses the upcoming concept of a Cybersecurity Requirements Specification (CRS) originating from the ISA TR 84.00.09 and the IEC 62443. The information currently stored within the attestation document already cover certain aspects of the proposed CRS, e.g. risk assessment results, coupled safety characteristics, or network segmentation based on zones and conduits.

The final aspect regarding a possible knowledge transfer aims at the publication of the proposed AAS submodel for automated security risk assessments at the IDTA[10]. There, a public track process is described which should be followed as a next step in alignment with the other participating working groups, such as the NAMUR working group for security engineering[11].

[9] https://www.iec.ch/dyn/www/f?p=103:30:::::FSP_ORG_ID:1250
[10] https://industrialdigitaltwin.org/en/content-hub/create-a-submodel
[11] https://industrialdigitaltwin.org/en/content-hub/submodels

Part VI.

Annex

14. Annex A - Swimlane Figures for the Information Collection

The original swimlanes for the first method step of information collection (see Section 6.2) are specified and visually documented using the MindManager[1] software due to the availability of a fitting license via the OWL University of Applied Sciences and Arts. The overall swimlane map is too big to be displayed here within the Annex in one piece. Therefore, the following eight figures on the subsequent pages show the contents in a rotated view to fit to the size limits.

[1] https://www.mindmanager.com/en

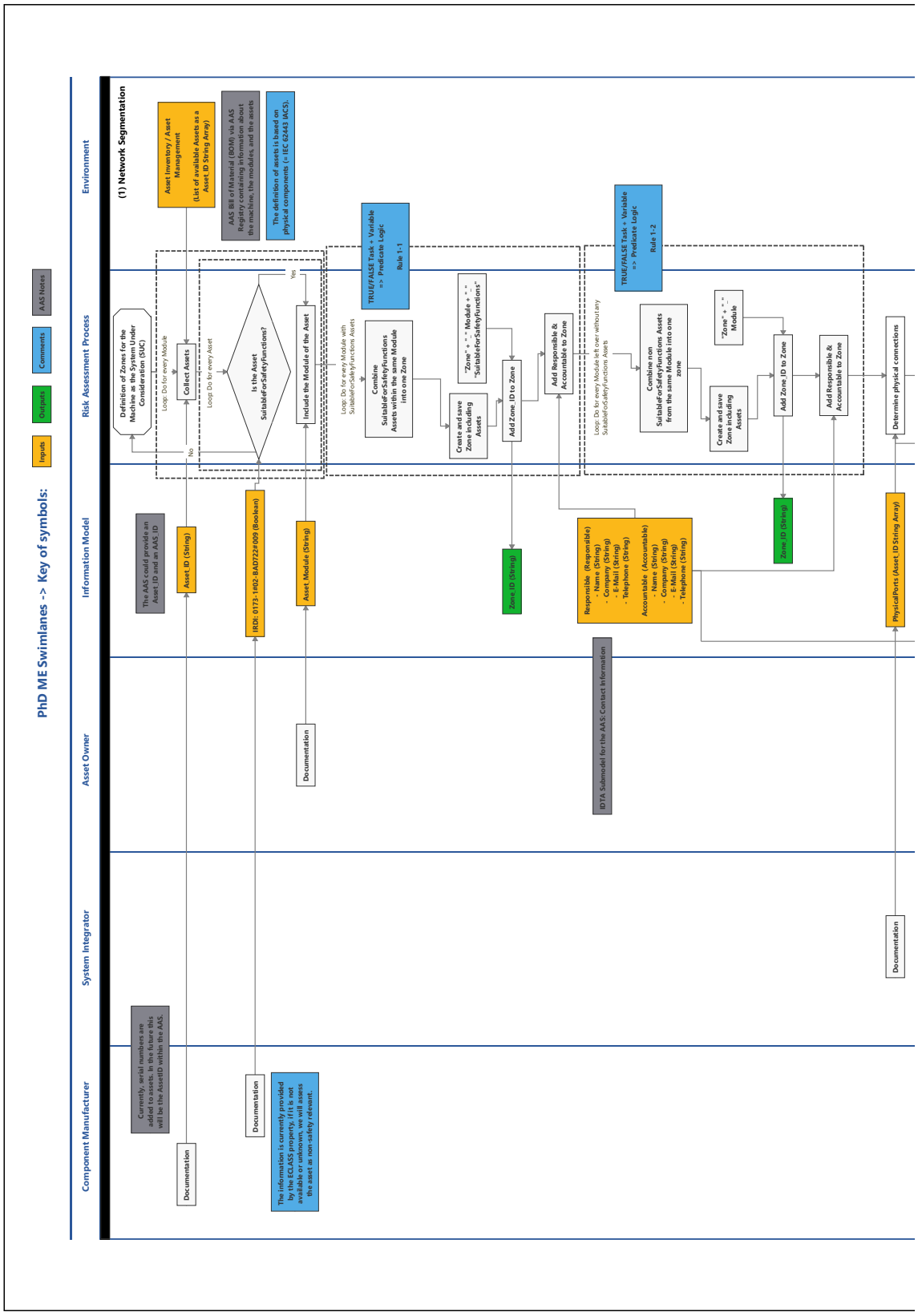

Figure 14.1: Detailed swimlanes of the security risk assessment process of this dissertation (1/8)

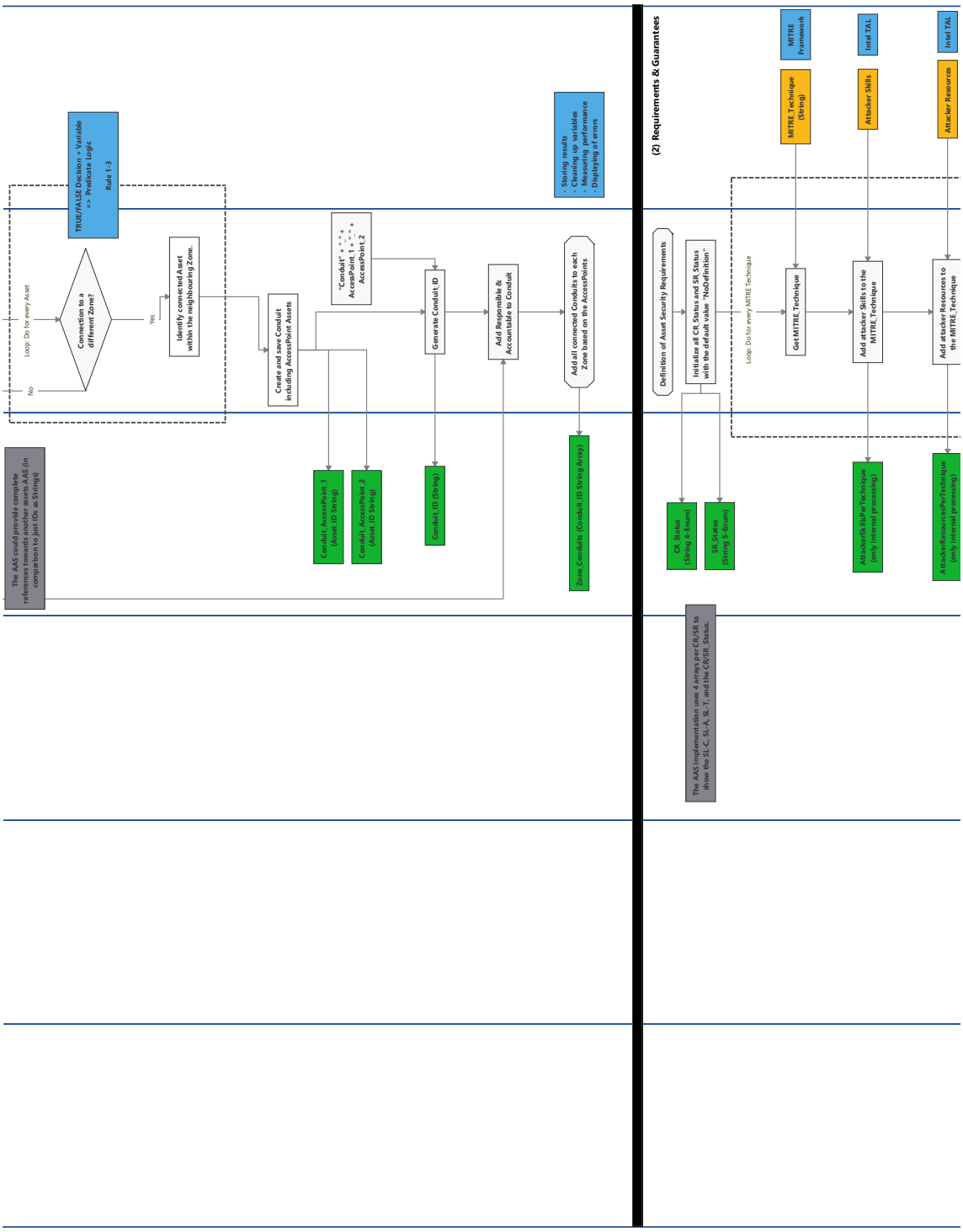

Figure 14.2: Detailed swimlanes of the security risk assessment process of this dissertation (2/8)

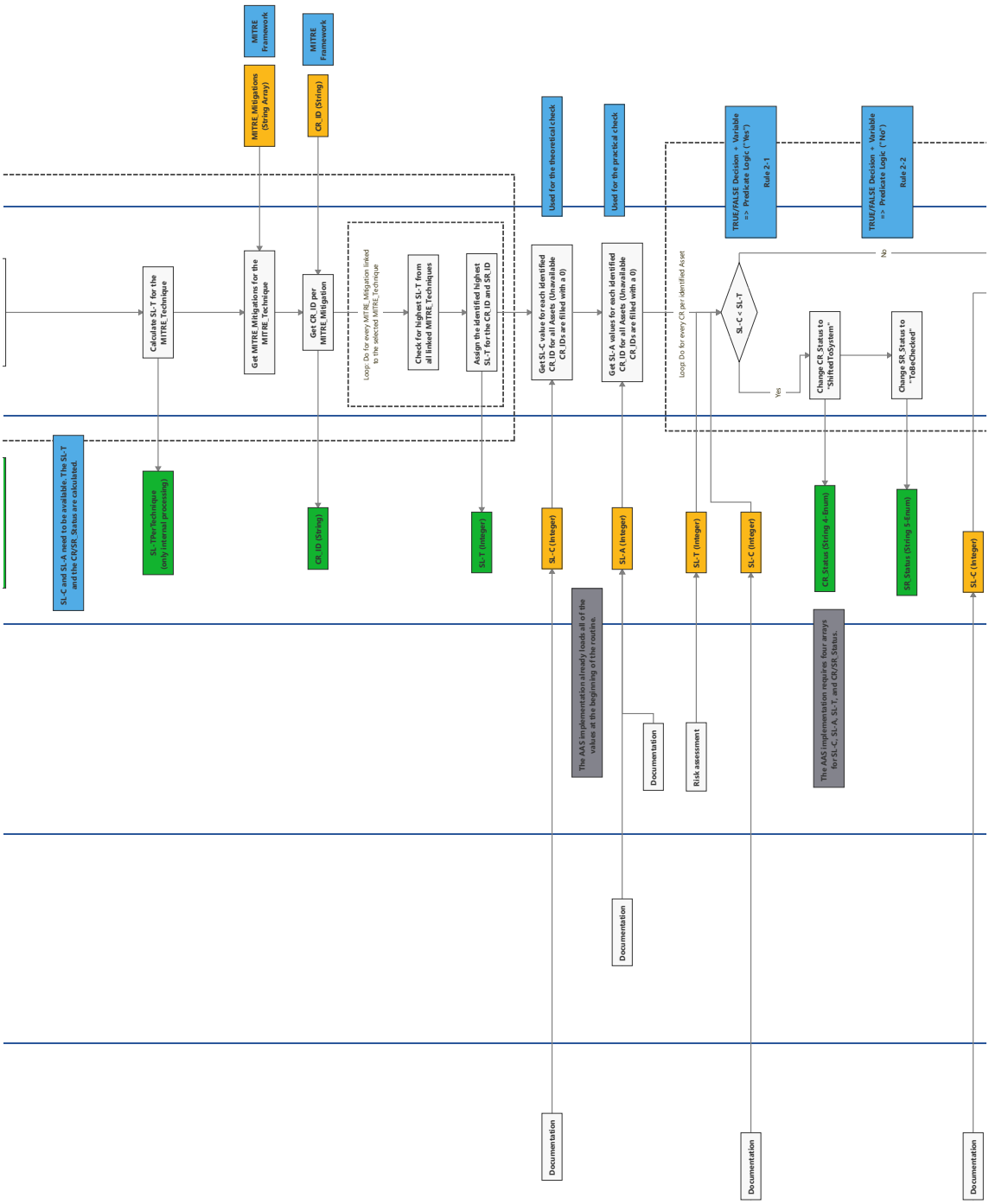

Figure 14.3: Detailed swimlanes of the security risk assessment process of this dissertation (3/8)

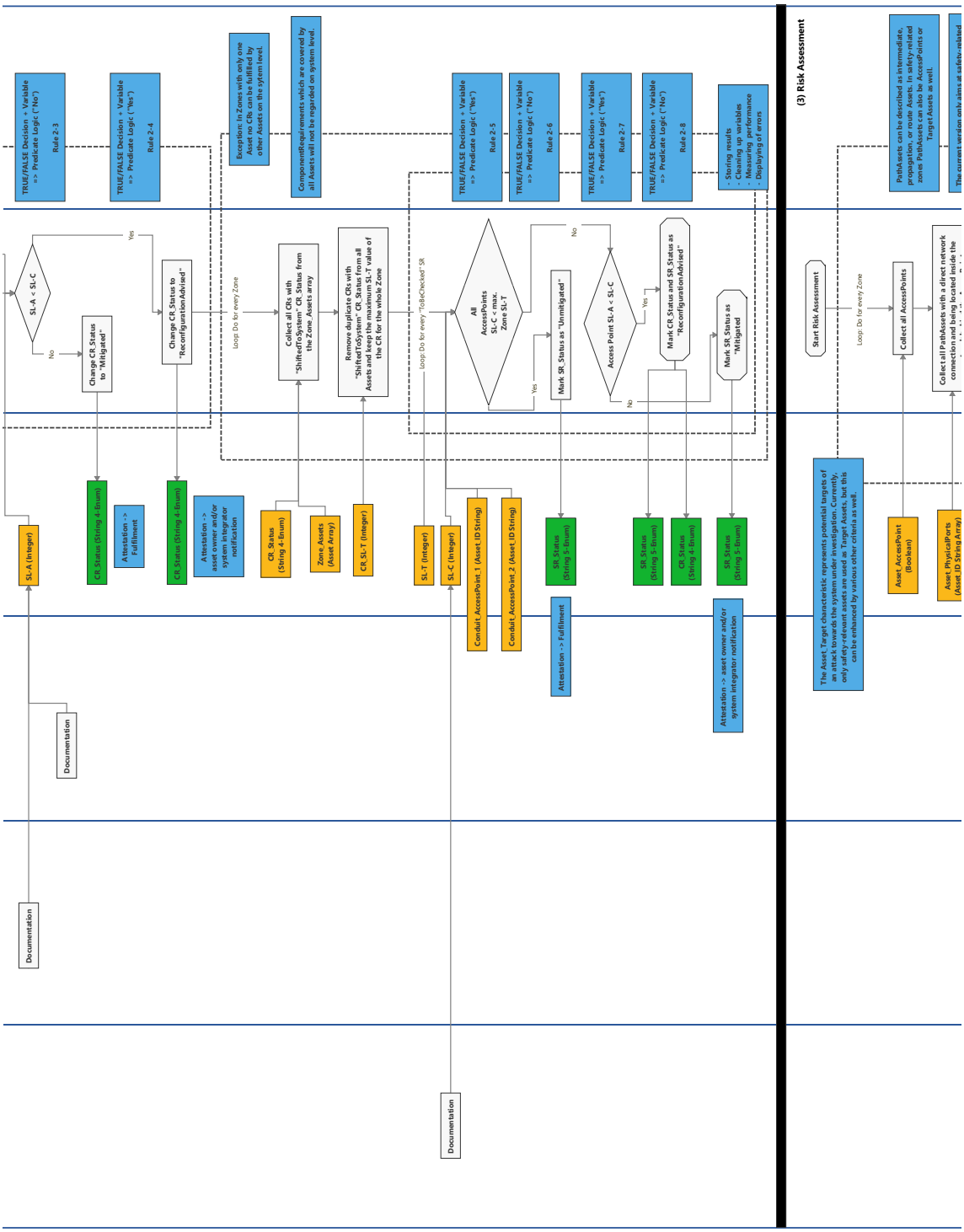

Figure 14.4: Detailed swimlanes of the security risk assessment process of this dissertation (4/8)

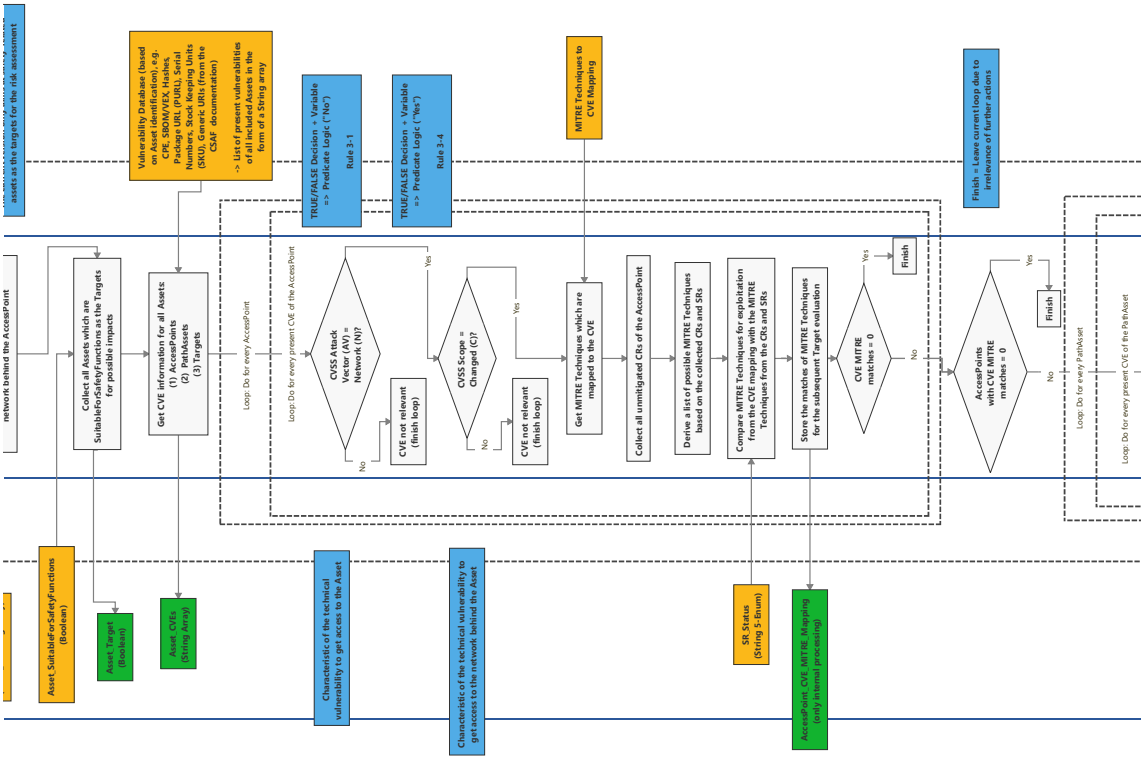

Figure 14.5: Detailed swimlanes of the security risk assessment process of this dissertation (5/8)

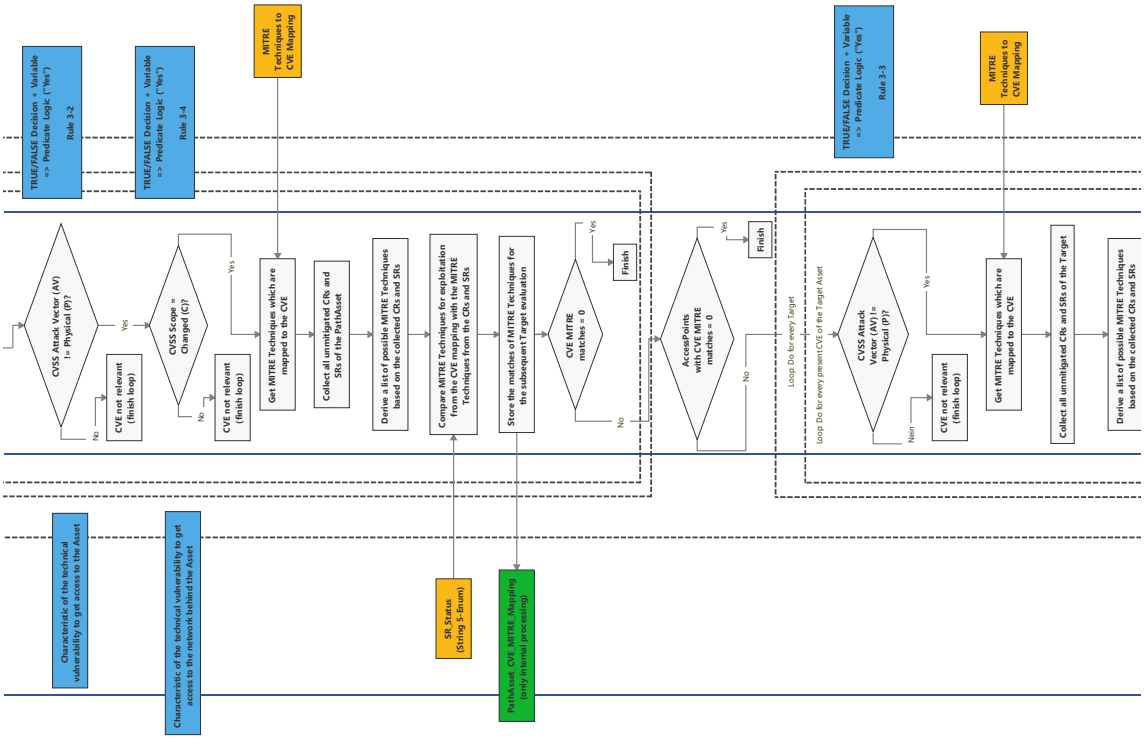

Figure 14.6: Detailed swimlanes of the security risk assessment process of this dissertation (6/8)

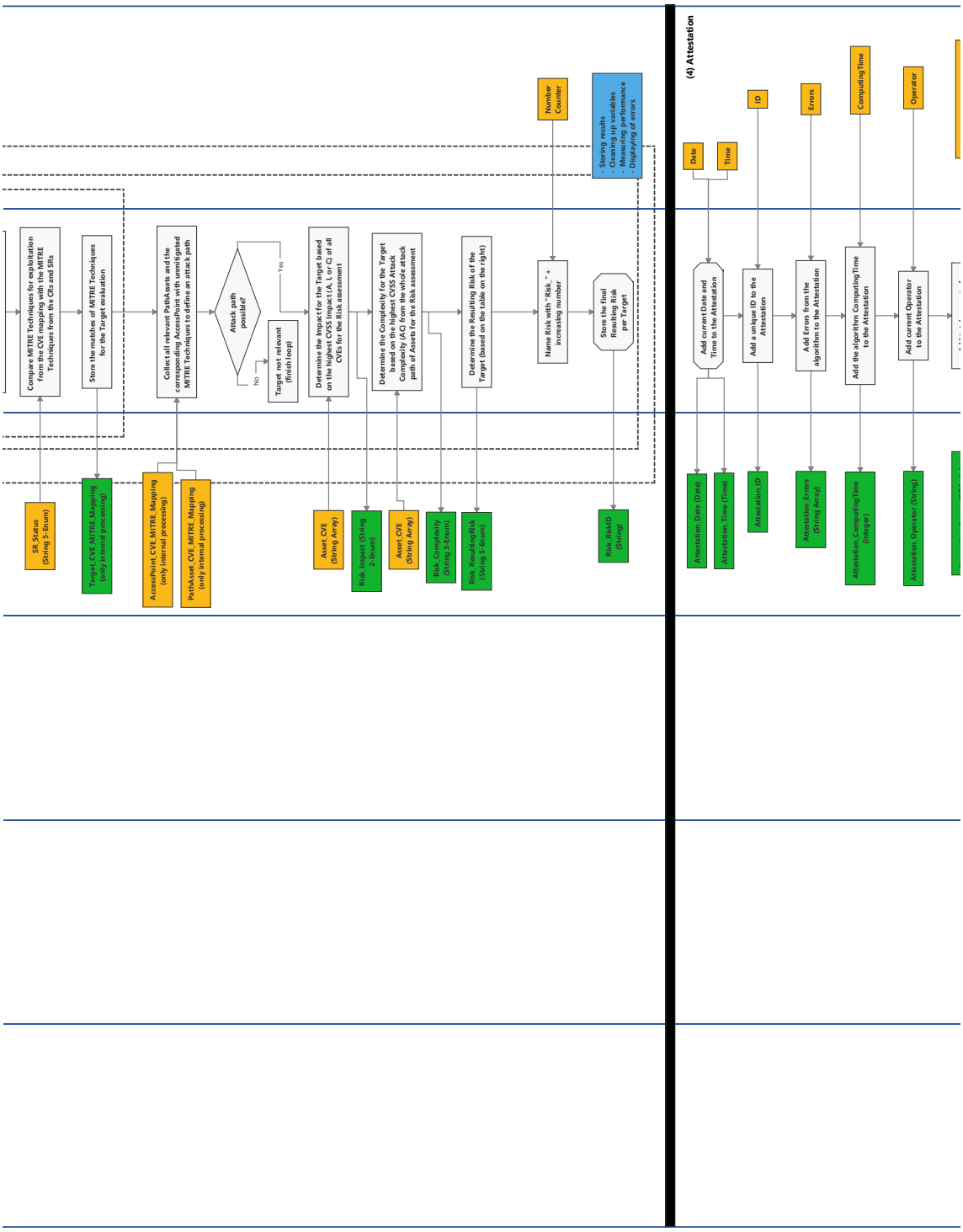

Figure 14.7: Detailed swimlanes of the security risk assessment process of this dissertation (7/8)

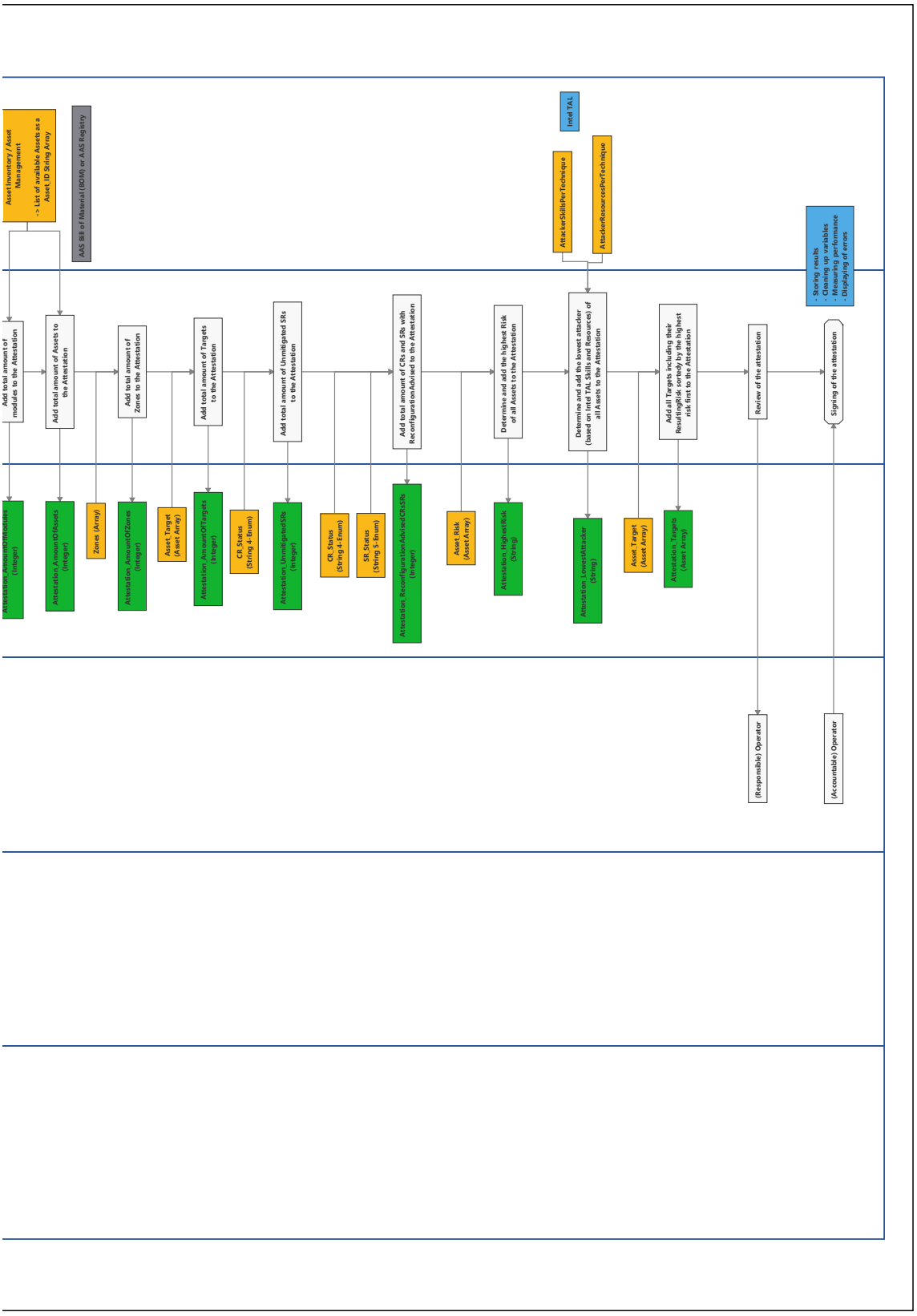

Figure 14.8: Detailed swimlanes of the security risk assessment process of this dissertation (8/8)

15. Annex B - Swimlane Example: SL-T Determination (ZCR 5.6)

This annex contains one explanatory example of the information collection and process modelling (excerpted from the authored and already published publication [94]) using the swimlane methodology regarding the determination of SL-T from the ZCR 5.6 defined within the IEC 62443-3-2 standard. This increases the comprehensibility of the underlying activities and the achieved results regarding the automation of security risk assessment processes.

Modelling Basics

The authors of [231] examine the three most popular attack models: Lockheed Martin's Cyber Kill Chain, the MITRE ATT&CK framework, and the Diamond Model. All the models offer different characteristics, advantages, and disadvantages. Annex E of the ISA-TR84.00.09 provides a comparison of Lockheed Martin's Cyber Kill Chain as a high-level model and the MITRE ATT&CK framework as a mid-level model. Lockheed Martin's Cyber Kill Chain is based on a seven phase approach to support defenders in understanding on how to break down an attack [231]. Nevertheless, it has several known security gaps and was not thoroughly updated since its creation in 2011 [231]. The Diamond Model was released in 2013 and offers support in investigating how an adversary exploits a capability of an infrastructure against a victim [231]. The usage of a mathematical framework makes it mostly too complex to be implemented and practically infeasible to use [231].

MITRE is an American non-profit organisation supporting research activities in various domains. In 2013, the MITRE Adversarial Tactics, Techniques, and Common Knowledge (ATT&CK) framework was originally released for enterprises and was since adapted for additional areas, e.g. mobile systems, cloud systems, or ICSs [232]. It can be used by both offensive and defensive sides of security teams for adversary emulation and defensive gap assessment by analysing, e.g. the formalised techniques with their associated mitigations as typical countermeasures [233]. Furthermore, it is accessible worldwide, widely accepted within the security domain, and already integrated into several technologies [231]. A recent study from the SANS Institute revealed that nearly 50% of respondents leverage the MITRE ATT&CK framework for ICS security within their organisation [234]. In addition, several businesses addressing ICS security are integrating the MITRE ATT&CK framework into their

OT monitoring solutions, such as Dragos[1] or Nozomi[2]. Therefore, this dissertation focuses on the MITRE ATT&CK framework for ICSs to be utilised for the SL-T determination demanded by the IEC 62443-3-2 standard.

In [235], the Intel TAL is originally presented. It describes twenty-two different threat agent archetypes (hostile or non-hostile), such as irrational individual, untrained employee, anarchist, terrorist, or government spy, with eight characteristics, including resources and skills. The typical usage and formalisation is based on threat identification, selection of possible threat agents, and assessment of threat agent skills and resources [236]. Alternatives regarding the Intel TAL are the ICS-CERT agent types [237] and CAPEC[3]. Both formalisations are not as extensive as the TAL and are typically used for other tasks, e.g. developer training or penetration testing.

The analysed state of the art shows that a variety of model-based security frameworks, taxonomies, and formalisations is available [176]. Nevertheless, a specific alignment of these security frameworks with the IEC 62443 standard concepts is missing right now within the research domain to fulfil the demands from the swimlanes.

Therefore, this dissertation proposes a specification of the ZCR 5.6 "Determine SL-T" security risk assessment step based on the identified security frameworks, especially the MITRE ATT&CK ICS techniques with associated mitigations and the Intel TAL characteristics regarding skills and resources.

Application Scenario

The SL-T determination method proposed in this work is explained with the help of a prominent example of a real-world incident regarding OT security. In 2017, the TRITON malware (also known as TRISIS or HatMan) affected a SIS, namely the Triconex controller from Schneider Electric [95]. This interrupted the operation at an oil and gas petrochemical facility in Saudi Arabia and was the first publicly reported incident in the context of SISs[4]. SISs are typically responsible for the safety monitoring of the facility's operation and designed to prevent damage towards humans, components, and the environment [95]. The authors of [95] show that the required resources for such an attack are only medium, and many different threat agents are theoretically capable of causing this kind of incident. Also, the MITRE ATT&CK framework provides a thorough analysis of the techniques used during the TRITON attack[5]. Therefore, it is important to include the already available and formalised knowledge about threat agents and their associated techniques into typical security risk assessments. Furthermore, the determination of the required

[1] https://www.dragos.com/mitre-attack-for-ics
[2] https://www.nozominetworks.com/blog/enhancing-threat-intelligence-with-the-mitre-attck-framework
[3] https://capec.mitre.org/about/attack_comparison.html
[4] https://attackevals.mitre-engenuity.org/ics/triton
[5] https://attack.mitre.org/software/S1009

SL-T for a system under consideration is vital to decide about mitigations and their necessity.

SL-T Determination Methodology

The TRITON attack is based on a sequence of techniques used by threat agents, which are already formalised within the MITRE ATT&CK ICS framework[6]. This dissertation uses the specific technique of "Remote System Discovery"[7] from the "Discovery" tactic as a representative and guiding example. In order to determine the SL-T, four steps need to be performed, which are further summarised within Figure 15.1 and described in a more detail afterwards.

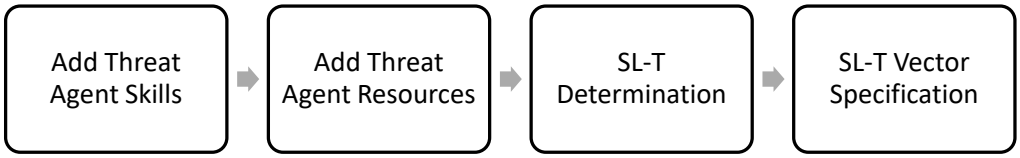

Figure 15.1: Overview of necessary steps for the SL-T determination

Add Threat Agent Skills: The Intel TAL provides skill levels and their definitions of the threat agents used for the proposed formalisation. This information can be used to further specify the "Remote System Discovery" technique in a manual preparation. In this case, it is decided by the author of this dissertation that the necessary skill level for the TRITON attack can be set to "Adept" based on the description of the incident mentioned before and due to the analysed medium level of required resources enhanced by the need for very specialised SIS know-how [95]. An "Adept" is described as an expert in technology and attack methods being able to apply existing attacks and create new ones to greatest advantage [235].

Add Threat Agent Resources: Moreover, the Intel TAL defines required resources for threat agents to perform certain attacks. In a second manual preparation step, the necessary resources for the TRITON attack can be defined as an intermediate step for the SL-T determination. The fitting classification is the resource "Organization" due to the fact that the Triton attack was analysed and categorised below a governmental or state-sponsored level [95]. An "Organization" is typically specified as a company-like structure with larger and better resources than those of a small team, usually operating in multiple countries globally and persisting on the long term [235].

SL-T Determination: In the next step, this dissertation proposes a matrix for the SL-T determination based on the previously defined threat agent skills and resources for the exemplary technique of "Remote System Discovery". Figure 15.2 is manually created as a preparatory step similar to the definition of typical risk matrices regarding style and contents, such as defined within the Annex B of the IEC

[6]https://attack.mitre.org/software/S1009
[7]https://attack.mitre.org/techniques/T0846

62443-3-2. The table is completely filled with values from left to right, increasing from zero to four, representing a gradual increase of the SL-T. Afterwards, unrealistic combinations of threat agent skills and resources are removed (represented as a crossed out values) due to minimum skill requirements for certain resources within the Intel TAL formalisation. Using this way of formalisation, the Intel TAL is able to represent typical threat agents and their associated characteristics in a realistic way. The provided example results in an SL-T of 3 based on the classifications of "Organization" resources and "Adept" skills as described above.

Resources / Skills	Individual	Club	Contest	Team	Organization	Government
None	0	~~1~~	~~1~~	~~2~~	~~2~~	~~2~~
Minimal	1	1	~~1~~	~~2~~	~~3~~	~~3~~
Operational	1	1	2	3	~~3~~	~~3~~
Adept	2	2	2	3	3	4

Figure 15.2: Mapping of Intel TAL skills and resources for the SL-T determination

SL-T Vector Specification: The usage of the Intel TAL and the MITRE ATT&CK framework provides the possibility to determine a certain SL-T value for a specific technique, in this example the technique "Remote System Discovery" as one step used for the TRITON incident. As mentioned before, there are various vector variants available within the IEC 62443 definitions. Here, a vector based on a set of the CRs and SRs is used to get a very detailed view on the SUC respecting the different characteristics defined for the component and system requirements. Each MITRE ATT&CK technique is mapped towards a fitting set of mitigations. In this example, the only available mitigation for the "Remote System Discovery" is the "Static Network Configuration" which demands static network configurations whenever possible[8]. The MITRE mitigations can then be mapped towards already available catalogues of mitigations from various standardisation documents, e.g. the NIST SP 800-53, the IEC 62443-3-3 SRs, or the IEC 62443-4-2 CRs. An older version of the MITRE ATT&CK ICS framework already provided this mapping by assigning one SR and CR for most of the mitigations, in our utilised version v13.1 the mapping was discarded. The newest version v14.1 of the MITRE ATT&CK ICS framework[9] now contains these information again. For the mitigations without an assigned SR or CR, it needs to be assigned in a manual way as an addition to the existing ones. The given example of the "Static Network Configuration" mitigation can be mapped towards the following requirements:

- NIST SP 800-53: CM-7 → "Least functionality"
- IEC 62443-3-3: SR 7.7 → "Least functionality"

[8]https://attack.mitre.org/mitigations/M0814
[9]https://attack.mitre.org/mitigations/ics

- IEC 62443-4-2: CR 7.7 → "Least functionality"

By doing so, it is possible to assign an SL-T that was determined based on the Intel TAL and the MITRE ATT&CK ICS technique to an IEC 62443-3-3 SR. As a result, the complete SL-T vector can be specified using existing security expert knowledge in a formalised way. Figure 15.3 sketches a possible SL-T vector based on the previously described results with the determined SL-T value of 3 for the CR/SR 7.7 highlighted in grey. The remaining CRs and SRs are shown here as well without values in order to illustrate the overall structure of the SL-T vector. However, they can be determined in the same way according to the given example presented before based on threat agent skills and resources.

FR 1		FR 2		FR 3		FR 4		FR 5		FR 6		FR 7	
CR/SR 1.1	-	CR/SR 2.1	-	CR/SR 3.1	-	CR/SR 4.1	-	CR/SR 5.1	-	CR/SR 6.1	-	CR/SR 7.1	-
...	-	...	-	...	-	...	-	...	-	CR/SR 6.2	-	...	-
...	-	...	-	...	-	CR/SR 4.3	-	...	-			CR/SR 7.7	3
CR/SR 1.14	-	CR/SR 2.13	-	CR/SR 3.14	-			CR/SR 5.4	-			CR/SR 7.8	-

Figure 15.3: Example of the determined SL-T vector in alignment to the IEC 62443 standard and the associated FRs

The proposed method for the determination of the SL-T vectors based on the Intel TAL and the MITRE ATT&CK ICS framework enables the elicitation of security expert knowledge during a one-time manual preparation phase to make it available for a subsequent automated processing during security risk assessments in conformance with the IEC 62443 standard. This permits the creation of consistent security risk assessment results in a transparent and repeatable manner. In addition, the proposed method for the ZCR 5.6 "Determine SL-T" improves the security of modular manufacturing systems and can potentially be integrated into automated security risk assessment processes. This is achieved by the possibility to clearly define a desired target state of the SUC. Hence, the target can be communicated to all involved stakeholders and mapped to available mitigation measures. Furthermore, the presented method enables the comparison of the determined SL-T with available metrics, such as the SL-C or the SL-A values of components.

16. Annex C - Manual Reference Security Risk Assessment

Table 16.1: Finding tracker of the manual reference security risk assessment for evaluation and result discussion

ID	Title	Complexity	Impact	Risk	Assets
NET-1	Insufficient segmentation of safety-related components	High	High	Medium	CabinetSwitch, ConveyorSwitch & LaserSwitch
VLN-1	Outdated Windows version	Medium	High	High	MES (within the public network)
VLN-2	Network switch vulnerable to Denial of Service (DoS)	High	High	Medium	CabinetSwitch
VLN-3	Safety PLCs vulnerable to remote code execution	High	High	Medium	CabinetSafetyPLC #1, #2 & #4
VLN-4	Network switches vulnerable to brute-force attacks	High	High	Medium	ConveyorSwitch & LaserSwitch
VLN-5	Laser controller vulnerable to unauthenticated manipulation	High	High	Medium	Laser Controller
ISMS-1	Lack of vulnerability and patch management	High	High	Medium	Organisational Finding

Additional descriptions for the findings:

- NET-1: The network is flat and contains all controllers regardless of their safety requirements. This allows an attacker who compromised any of the components, including the remote access client, to do lateral movement to the safety PLCs. Without a firewall between safety zones and all external components, a full compromise is possible.
- VLN-1: The Windows machine hosting the MES is an old version that is not tracked, probably has unidentified vulnerabilities, and does not have a patch management process implemented. An attacker can take over control of the machine and manipulate the production process, possibly comprising human safety.

- VLN-2: The cabinet switch in use has multiple vulnerabilities (CVE-2020-6994, CVE-2020-9307), that allow an attacker to reconfigure it and/or to cause a disruption in the network service, which can impact production. To exploit this, an attacker needs to be able to connect to the cabinet switch from the Local Area Network (LAN) or using the remote access.
- VLN-3: The safety PLCs can be fully taken over by an attacker on the network, including a manipulation of production and violation of the safety guarantees (CVE-2020-12519). They can be forced to provide incorrect safety signals, possibly leading to human injury.
- VLN-4: The conveyor and laser switches expose administrative interfaces that lack brute-force protection (CVE-2019-16670; CVE-2018-5469). A determined attacker can attempt to try a large number of passwords and eventually gain access to the network switches. This can lead to a DoS attack on the production line.
- VLN-5: The laser controller web interface can be accessed without knowing the administrator credentials (CVE-2016-8371). Furthermore, the variables can be manipulated without a login over the network (CVE-2016-8380). This can be used to manipulate / start the laser without ensuring that all safety preconditions are met, potentially causing human injury.
- ISMS-1: There is no process in place to ensure that vulnerabilities in the components of the production line are identified and patched or mitigated.

17. Annex D - List of Publications

The following list summarises the publications, which are important for this dissertation, in alignment with Figure 1.6 from Section 1.4. It contains the title, year, conference, authorship, and reference of each publication:

Automation 2017 (first author): Security Concept for a Cloud-based Automation Service [8]

ETFA 2017 (first author): Automatic Mapping of Cyber Security Requirements to support Network Slicing in Software-Defined Networks [181]

WFCS 2018 (first author): Modelling and Automatic Mapping of Cyber Security Requirements for Industrial Applications: Survey, Problem Exposition, and Research Focus [238]

Echtzeit 2018 (first author): Automatische Evaluierung von Anforderungen bezüglich der Informationssicherheit für das zukünftige industrielle Netzwerkmanagement [239]

KommA 2018 (co-author): Modelling Security Requirements and Controls for an Automated Deployment of Industrial IT Systems [135]

at Schwerpunktheft 2018 (first author): Quality-of-Service Monitoring of Hybrid Industrial Communication Networks [23]

IECON 2019 (first author): Survey of Security Standards for an automated Industrie 4.0 compatible Manufacturing [39]

NetSys 2019 (first author): Automated Processing of Security Requirements and Controls for a common Industrie 4.0 Use Case [121]

ETFA 2019 (first author): Secure and Flexible Deployment of Industrial Applications inside Cloud-Based Environments [9]

ETFA 2020 (first author): Towards Automated Security Evaluation within the Industrial Reference Architecture [229]

KommA 2020 (first author): Automatische Bewertung und Überwachung von Safety & Security Eigenschaften: Strukturierung und Ausblick [41]

e&i 2021 (first author): Alignment of safety and security risk assessments for modular production systems [31]

WFCS 2022 (first author): Investigation of Resource Constraints for the Automation of Industrial Security Risk Assessments [53]

ETFA 2022 (co-author): Towards an Asset Administration Shell Integrity Verification Scheme [214]

KommA 2022 (co-author): An Asset Administration Shell Version Control to Enforce Integrity Protection [210]

EKA 2022 (first author): Towards Automated Risk Assessments for Modular Industrial Automation and Control Systems - State of the Art Survey and Information Model Proposal [56]

at-Schwerpunktheft 2023 (first author): Towards Automated Risk Assessments for Modular Manufacturing Systems - Process Analysis and Information Model Proposal [57]

INDIN 2023 (first author): Determining the Target Security Level for Automated Security Risk Assessments [94]

GitHub Repository[1] (first published in 2023)

ETFA 2023 (first author): Evaluation Concept for Prototypical Implementation towards Automated Security Risk Assessments [227]

WFCS 2024 (first author): Requirements Analysis for the Evaluation of Automated Security Risk Assessments [179]

ETFA 2024 (first author): Evaluation of an Automated Security Risk Assessment based on a Manual Reference (submitted and still under review)

KommA 2024 (co-author): Automatische Bewertung und Überwachung von Safety- & Security-Eigenschaften: Konzept, Informationsmodelle und Herausforderungen (submitted and still under review)

Automation 2024 (co-author): Evolution der IT/OT-Security durch modulare Anlagenkonzepte (submitted and still under review)

NAMUR NE 193 (co-author): Ein Informationsmodell für das Automation Security Engineering (submitted and still under review)

TÜV Rheinland Symposium 2024 (co-author): Automatic hazard and risk assessment for a flexible production (submitted and still under review)

[1] https://github.com/auto-s2/security-risk-assessment

Part VII.

Directories

Abbreviations

AAS	Asset Administration Shell
ACL	Access Control List
ACT	Attack Countermeasure Tree
AG	Attack Graph
AI	Artificial Intelligence
AML	Automation Markup Language
AMS	Adjustable Manufacturing System
API	Application Programming Interface
ATA	Attack Tree Analysis
BAG	Bayesian Attack Graph
BBN	Bayesian Belief Network
BDMP	Boolean logic Driven Markov Processes
BN	Bayesian Network
BOM	Bill of Material
BPMN	Business Process Model Notation
BSI	Bundesamt für Sicherheit in der Informationstechnik
CAEX	Computer Aided Engineering Exchange
CAPEC	Common Attack Pattern Enumeration and Classification
CAPEX	Capital Expenditure
CDD	Common Data Dictionary
CE	Conformité Européenne
CERT	Computer Emergency Response Team
CIP	Critical Infrastructure Protection
CMO	Current Mode of Operation
COTS	Commercial off-the-Shelf
CPE	Common Platform Enumeration
CPPS	Cyber-Physical Production System
CPS	Cyber-Physical System

CR	Component Requirement
CRS	Cybersecurity Requirements Specification
CSAF	Common Security Advisory Framework
CSET	Cyber Security Evaluation Tool
CSF	Cybersecurity Framework
CSMS	Cyber Security Management System
CTMC	Continuous Time Markov Chain
CVE	Common Vulnerability Enumeration
CVSS	Common Vulnerability Scoring System
CWE	Common Weakness Enumeration
CWSS	Common Weakness Scoring System
CySeMoL	Cyber Security Modelling Language
DAG	Directed Acyclic Graph
DCRA	Detailed Cybersecurity Risk Assessment
DFD	Data Flow Diagram
DHS	Department of Homeland Security
DMS	Dedicated Manufacturing System
DMZ	Demilitarized Zone
DoS	Denial of Service
DT	Digital Twin
DTMC	Discrete Time Markov Chain
DREAD	Damage, Reproducibility, Exploitability, Affected Users, and Discoverability
EN	European Norm
ERP	Enterprise Resource Planning
ETA	Event Tree Analysis
EU	European Union
FCAPS	Fault, Configuration, Accounting, Performance and Security
FEA	Functional Equipment Assembly
FMEA	Failure Modes and Effects Analysis
FMO	Future Mode of Operation
FMS	Flexible Manufacturing System
FR	Foundational Requirement

FTA	Fault Tree Analysis
GDPR	General Data Protection Regulation
GUI	Graphical User Interface
GRC	Governance, Risk Management, and Compliance
HMI	Human-Machine Interface
HTTP	Hypertext Transfer Protocol
IACS	Industrial Automation and Control System
ICRA	Initial Cybersecurity Risk Assessment
ICS	Industrial Control System
ICT	Information and Communication Technology
IDS	Intrusion Detection System
IDTA	Industrial Digital Twin Association
IEC	International Electrotechnical Commission
I4.0	Industrie 4.0
II	Industrial Internet
IPS	Intrusion Prevention System
IRDI	International Registration Data Identifier
IRI	Internationalized Resource Identifier
ISA	International Society of Automation
ISMS	Information Security Management System
ISO	International Organization for Standardization
IT	Information Technology
JSON	JavaScript Object Notation
KG	Knowledge Graph
KPI	Key Performance Indicator
LAN	Local Area Network
LARS ICS	Light and Right Security ICS
LOA	Level of Autonomy
LOK	Level of Knowledge
MBSE	Model-Based System Engineering
MES	Manufacturing Execution System
ML	Machine Learning
MTP	Module Type Package

MQTT	Message Queuing Telemetry Transport
NASA	National Aeronautics and Space Administration
NETCONF	The Network Configuration Protocol
NERC	North American Electric Reliability Corporation
NIST	National Institute of Standards and Technology
NVD	National Vulnerability Database
OAHP	Optimized Analytic Hierarchy Process
OODA	Observe, Orient, Decide, and Act
OPC	Open Platform Communications
OPC UA	Open Platform Communications Unified Architecture
OPEX	Operational Expenditure
OS	Operating System
OT	Operational Technology
OWASP	Open Worldwide Application Security Project
OWL	Web Ontology Language
PAN	Production Asset Network
PDCA	Plan, Do, Check and Act
PDF	Portable Document Format
PEA	Process Equipment Assembly
PERA	Purdue Enterprise Reference Architecture
PHA	Process Hazard Analysis
PI	PROFIBUS & PROFINET International
PI4.0	Plattform Industrie 4.0
PL	Performance Level
PLC	Programmable Logic Controller
POL	Process Orchestration Layer
PPR	Product-Process-Resource
PRA	Probabilistic Risk Assessment
PSA	Probabilistic Safety Assessment
QRA	Quantitative Risk Assessment
RAMI4.0	Reference Architecture Model Industrie 4.0
RC	Rule Class
RDF	Resource Description Framework

RDFS	Resource Description Framework Schema
REST	Representational State Transfer
RM	Risk Management
RMS	Reconfigurable Manufacturing System
SBOM	Software Bill of Material
SCADA	Supervisory Control and Data Acquisition
SCAP	Security Content Automation Protocol
SecDSVL	Security Domain Specific Visual Language
SHACL	Shapes Constraint Language
SL	Security Level
SIEM	Security Information and Event Management
SIL	Safety Integrity Level
SIS	Safety Instrumented System
SME	Small and Medium-sized Enterprise
SP	Security Program
SPARQL	SPARQL Protocol and RDF Query Language
SR	System Requirement
SRE	Security Requirements Engineering
STRIDE	Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privileges
SUC	System Under Consideration
SWRL	Semantic Web Rule Language
TAL	Threat Agency Library
TARA	Threat And Risk Analysis
TCP	Transmission Control Protocol
TISAX	Trusted Information Security Assessment Exchange
TLS	Transport Layer Security
TR	Technical Report
TRL	Technology Readiness Level
UDP	User Datagram Protocol
UML	Unified Modelling Language
VDE	Verband der Elektrotechnik Elektronik Informationstechnik
VDI	Verein deutscher Ingenieure

VEX	Vulnerability Exploitability eXchange
VLMC	Variable Length Markov Chain
VPN	Virtual Private Network
XML	Extensible Markup Language
YANG	Yet Another Next Generation
ZCR	Zone and Conduit Requirement

List of Figures

List of Tables

Bibliography

[1] A. Lüder, H. Steininger, and D. Goltz, *Quo vadis Automation? Trends für das Engineering von Automatisierungssystemen*, 17. Fachtagung EKA – Entwurf komplexer Automatisierungssysteme, Magdeburg, Germany, 2022.

[2] S. Eggers and K. Le Blanc, *Survey of Cyber Risk Analysis Techniques for Use in the Nuclear Industry*, Progress in Nuclear Energy 140, Elsevier, 2021.

[3] D. Etz, P. Denzler, T. Frühwirth, and W. Kastner, *Functional Safety Use Cases in the Context of Reconfigurable Manufacturing Systems*, 27th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Stuttgart, Germany, 2022.

[4] M. Mytych, L. Wisniewski, and J. Jasperneite, *Potential of SDN/NFV Network Management Concepts with Regards to Industrial Automation*, 8. Jahreskolloquium Kommunikation in der Automation (KommA), Magdeburg, Germany, 2017.

[5] J.-Q. Li, F. R. Yu, G. Deng, C. Luo, Z. Ming, and Q. Yan, *Industrial Internet: A Survey on the Enabling Technologies, Applications, and Challenges*, IEEE Communications Surveys & Tutorials, Volume: 19, Issue: 3, 2017.

[6] A. Klose *et al.*, *Building Blocks for Flexible Functional Safety in Discrete Manufacturing and Process Industries*, 26th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Västerås, Sweden, 2021.

[7] M. Ehrlich, H. Trsek, L. Wisniewski, and J. Jasperneite, *Modelling and Automatic Mapping of Cyber Security Requirements for Industrial Applications: Survey, Problem Exposition, and Research Focus*, 14th IEEE International Workshop on Factory Communication Systems (WFCS), Imperia, Italy, 2018.

[8] M. Ehrlich, L. Wisniewski, H. Trsek, D. Mahrenholz, and J. Jasperneite, *Automatic Mapping of Cyber Security Requirements to support Network Slicing in Software-Defined Networks*, 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol, Cyprus, 2017.

[9] M. Ehrlich, H. Trsek, M. Gergeleit, J. Paffrath, K. Simkin, and J. Jasperneite, *Secure and Flexible Deployment of Industrial Applications inside Cloud-Based Environments*, 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain, 2019.

[10] R. Paes, D. C. Mazur, B. K. Venné, and J. Ostrzenski, *A Guide to Securing Industrial Control Networks: Integrating IT and OT Systems*, IEEE Industry Applications Magazine, Volume: 26, Issue: 2, 2020.

[11] M. Ehrlich, H. Trsek, and J. Jasperneite, *Automatische Evaluierung von Anforderungen bezüglich der Informationssicherheit für das zukünftige industrielle Netzwerkmanagement*, Echtzeit und Sicherheit, Boppard, Germany, 2018.

[12] M. Wollschlaeger, T. Sauter, and J. Jasperneite, *The Future of Industrial Communication: Automation Networks in the Era of the Internet of Things and Industry 4.0*, IEEE Industrial Electronics Magazine, March, 2017.

[13] M. Ehrlich, L. Wisniewski, and J. Jasperneite, *State of the Art and Future Applications of Industrial Wireless Sensor Networks*, 7. Jahreskolloquium Kommunikation in der Automation (KommA), Lemgo, Germany, 2016.

[14] S. S. P. Olaya, A. Winkel, M. Ehrlich, M. Majumder, A. Schupp, and M. Wollschlaeger, *CANopen Flying Master over TSN*, 11. Jahreskolloquium Kommunikation in der Automation (KommA), Lemgo, Germany, 2020.

[15] S. Scanzio, L. Wisniewski, and P. Gaj, *Heterogeneous and Dependable Networks in Industry: A Survey*, Computers in Industry, Volume 125, 2021.

[16] A. Neumann, L. Wisniewski, R. S. Ganesan, P. Rost, and J. Jasperneite, *Towards Integration of Industrial Ethernet with 5G Mobile Networks*, 14th IEEE International Workshop on Factory Communication Systems (WFCS), Imperia, Italy, 2018.

[17] J. Jasperneite, T. Sauter, and M. Wollschlaeger, *Why We Need Automation Models: Handling Complexity in Industry 4.0 and the Internet of Things*, IEEE Industrial Electronics Magazine, March, 2020.

[18] M. Glawe and A. Fay, *Wissensbasiertes Engineering automatisierter Anlagen unter Verwendung von AutomationML und OWL*, at – Automatisierungstechnik, Volume: 64, Issue: 3, 2016.

[19] M. Ehrlich, L. Wisniewski, and J. Jasperneite, *Usage of Retrofitting for Migration of Industrial Production Lines to Industry 4.0*, 6. Jahreskolloquium Kommunikation in der Automation (KommA), Magdeburg, Germany, 2015.

[20] M. Cheminod, L. Durante, and A. Valenzano, *Review of Security Issues in Industrial Networks*, IEEE Transactions on Industrial Informatics, Volume: 9, Issue: 1, 2013.

[21] A. Neumann, M. Ehrlich, L. Wisniewski, and J. Jasperneite, *Towards Monitoring of Hybrid Industrial Networks*, 13th IEEE International Workshop on Factory Communication Systems (WFCS), Trondheim, Norway, 2017.

[22] M. Ehrlich, H. Trsek, D. Lang, L. Wisniewski, V. Wendt, and J. Jasperneite, *Security Concept for a Cloud-based Automation Service*, VDI-Kongress Automation, Baden-Baden, Germany, 2017.

[23] M. Ehrlich, A. Neumann, A. Biendarra, and J. Jasperneite, *Quality-of-Service Monitoring of Hybrid Industrial Communication Networks*, at Automatisierungstechnik Schwerpunktheft: Funk in der Automation, De Gruyter, 2018.

[24] A. Frimpong *et al.*, *Controller of Controllers Architecture for Management of Heterogeneous Industrial Networks*, 16th IEEE International Workshop on Factory Communication Systems (WFCS), Porto, Portugal, 2020.

[25] L. Dürkop, J. Jasperneite, and A. Fay, *An Analysis of Real-Time Ethernets with regard to their Automatic Configuration*, 11th IEEE International Workshop on Factory Communication Systems (WFCS), Palma de Mallorca, Spain, 2015.

[26] D. Lang, M. Friesen, M. Ehrlich, L. Wisniewski, and J. Jasperneite, *Pursuing the Vision of Industrie 4.0 - Secure Plug and Produce by Means of the Asset Administration Shell and Blockchain Technology*, IEEE 16th International Conference on Industrial Informatics (INDIN), Porto, Portugal, 2018.

[27] S. K. Panda, L. Wisniewski, M. Ehrlich, M. Majumder, and J. Jasperneite, *Plug and Play Retrofitting Approach for Data Integration to the Cloud*, 16th IEEE International Workshop on Factory Communication Systems (WFCS), Porto, Portugal, 2020.

[28] S. Fluchs, R. Drath, and A. Fay, *A Security Decision Base: How to Prepare Security by Design Decisions for Industrial Control Systems*, 17. Fachtagung EKA – Entwurf komplexer Automatisierungssysteme, Magdeburg, Germany, 2022.

[29] S. Kriaa, L. Pièetre-Cambacédès, M. Bouissou, and Y. Halgand, *A Survey of Approaches Combining Safety and Security for Industrial Control Systems*, Reliability Engineering and System Safety 139, 2015.

[30] B. Schneier, *Click Here to Kill Everybody: Security and Survival in a Hyper-Connected World*, W. W. Norton and Company, 2018.

[31] M. Ehrlich *et al.*, *Alignment of Safety and Security Risk Assessments for Modular Production Systems*, e&i Elektrotechnik und Informationstechnik, 2021.

[32] E. Tastan, S. Fluchs, and R. Drath, *Security Engineering with AutomationML: A Methodology for Modeling Security Decisions, Goals, Risks, and Requirements*, VDI-Kongress Automation, Baden-Baden, Germany, 2023.

[33] M. Carl and K. Gondlach, *Sicherheit 2027: Konformitätsbewertung in einer digitalisierten und adaptiven Welt*, Trendstudie des 2b AHEAD ThinkTanks, 2017.

[34] Y. Koren *et al.*, *Reconfigurable Manufacturing Systems*, CIRP Annals - Manufacturing Technology, 1999.

[35] Y. Y. Yusuf, M. Sarhadi, and A. Gunasekaran, *Agile Manufacturing: The Drivers, Concepts, and Attributes*, International Journal of Production Economics, 1999.

[36] Plattform Industrie 4.0, *Diskussionspapier: Security in der Verwaltungsschale*, German Federal Ministry for Economic Affairs and Energy (BMWi), 2017.

[37] G. Kavallieratos, S. Katsikas, and V. Gkioulos, *Cybersecurity and Safety Co-Engineering of Cyberphysical Systems: A Comprehensive Survey*, Future Internet 12, No. 4, 2020.

[38] H. Trsek, *Isochronous Wireless Network for Real-time Communication in Industrial Automation*, Dissertation, Fakultät für Informatik der Otto-von-Guericke-Universität Magdeburg, 2015.

[39] M. Ehrlich, M. Gergeleit, K. Simkin, and H. Trsek, *Automated Processing of Security Requirements and Controls for a common Industrie 4.0 Use Case*, International Conference on Networked Systems Workshop - Advanced Communication Networks for Industrial Applications, Garching, Germany, 2019.

[40] J. de La Motte and P. Sieber, *Safety goes digital: Digitalisierung der funktionalen Sicherheit mit Mehrwert*, atp, Funktionale Sicherheit, Ausgabe 10, 2022.

[41] M. Ehrlich *et al.*, *Automatische Bewertung und Überwachung von Safety & Security Eigenschaften: Strukturierung und Ausblick*, 11. Jahreskolloquium Kommunikation in der Automation (KommA), Lemgo, Germany, 2020.

[42] G. Lykou, A. Anagnostopoulou, G. Stergiopoulos, and D. Gritzalis, *Cybersecurity Self-assessment Tools: Evaluating the Importance for Securing Industrial Control Systems in Critical Infrastructures*, 13th International Conference on Critical Information Infrastructures Security (CRITIS), Kaunas, Lithuania, 2019.

[43] D. Etz, T. Frühwirth, and W. Kastner, *Flexible Safety Systems for Smart Manufacturing*, 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria, 2020.

[44] C. Tebbe, *Durchgängiges Wissensmanagement von OT-Security-Wissen im Lebensweg von Produktionsanlagen*, Dissertation, Fakultät für Maschinenbau der Helmut-Schmidt-Universität/Universität der Bundeswehr, Hamburg, 2021.

[45] N. Maunero, F. de Rosa, and P. Prinetto, *Towards Cybersecurity Risk Assessment Automation: An Ontological Approach*, IEEE International Symposium on Dependable, Autonomic and Secure Computing (DASC), Abu Dhabi, United Arab Emirates, 2023.

[46] C. Tebbe, M. Glawe, A. Scholz, K.-H. Niemann, A. Fay, and J. Dittgen, *Wissensbasierte Sicherheitsanalyse in der Automation*, atp-edition, Automatisierungstechnische Praxis, Issue: 4, 2015.

[47] S. Fluchs, R. Drath, and A. Fay, *Evaluation of Visual Notations as a Basis for ICS Security Design Decisions*, IEEE Access, Volume: 11, 2023.

[48] A. Shameli-Sendi, R. Aghababaei-Barzegar, and M. Cheriet, *Taxonomy of Information Security Risk Assessment (ISRA)*, Journal of Computers and Security, Elsevier, 2016.

[49] F. de Franco Rosa, R. Bonacin, and M. Jino, *The Security Assessment Domain: A Survey of Taxonomies and Ontologies*, PRJ04.35 – H1 14/01088 TRT0049417, Renato Archer Information Technology Center (CTI), 2017.

[50] B. Czybik, A. Oswald, M. Ehrlich, and H. Trsek, *Eine differenzierte Betrachtung ganzheitlicher Selbstorganisation im Rahmen von Industrie 4.0*, VDI-Kongress Automation, Baden-Baden, Germany, 2020.

[51] M. D. Aime, A. Atzeni, and P. C. Pomi, *AMBRA: Automated Model-Based Risk Analysis*, 2007.

[52] A. Bamberg, *Interview: Wir verdrahten nicht, wir starten.* atp, Funktionale Sicherheit, Ausgabe 10, 2022.

[53] M. Ehrlich, G. Lukas, H. Trsek, J. Jasperneite, and C. Diedrich, *Investigation of Resource Constraints for the Automation of Industrial Security Risk Assessments*, 18th IEEE International Workshop on Factory Communication Systems (WFCS), Pavia, Italy, 2022.

[54] M. Hamdi and N. Boudriga, *Algebraic Specification of Network Security Risk Management*, ACM Workshop on Formal Methods in Security Engineering, Washington, USA, 2003.

[55] M. Eckhart, A. Ekelhart, and E. Weippl, *Enhancing Cyber Situational Awareness for Cyber-Physical Systems through Digital Twins*, 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain, 2019.

[56] M. Ehrlich, A. Bröring, C. Diedrich, and J. Jasperneite, *Towards Automated Risk Assessments for Modular Industrial Automation and Control Systems: State of the Art Survey and Information Model Proposal*, 17. Fachtagung EKA – Entwurf komplexer Automatisierungssysteme, Magdeburg, Germany, 2022.

[57] M. Ehrlich, A. Bröring, C. Diedrich, and J. Jasperneite, *Towards Automated Risk Assessments for Modular Manufacturing Systems: Process Analysis and Information Model Proposal*, at – Automatisierungstechnik, Volume: 71, Issue: 6, 2023.

[58] S. Russel and P. Norvig, *Artifical Intelligence: A Modern Approach*, Fourth Edition, Pearson Series In Artificial Intelligence, 2021.

[59] M. Schumacher and U. Roedig, *Security Engineering with Patterns*, 8th Annual Conference on the Pattern Languages of Programs (PLoP), Urbana, USA, 2001.

[60] Z. A. Styczynski, K. Rudion, and A. Naumann, *Einführung in Expertensysteme: Grundlagen, Anwendungen und Beispiele aus der elektrischen Energieversorgung*, 1. Edition, Springer Vieweg, 2017.

[61] H. Bonnin and O. Flebus, *Data & Safety: Challenges and Opportunities*, 5th International Workshop on Critical Automotive Applications: Robustness & Safety, Naples, Italy, 2019.

[62] T. Aven, P. Baraldi, R. Flage, and E. Zio, *Uncertainty in Risk Assessment: The Representation and Treatment of Uncertainties by Probabilistic and Non-Probabilistic Methods*, John Wiley & Sons Ltd., 2014.

[63] Plattform Industrie 4.0, *Umgang mit Sicherheitsrisiken industrieller Anwendungen durch mangelnde Erklärbarkeit von KI-Ergebnissen*, German Federal Ministry for Economic Affairs and Energy (BMWi), 2019.

[64] Y. Cherdantseva *et al.*, *A Review of Cyber Security Risk Assessment Methods for SCADA Systems*, Computers & Security, Issue: 56, 2016.

[65] A. Cook, R. Smith, L. Maglaras, and H. Janicke, *Measuring the Risk of Cyber Attack in Industrial Control System*, 4th International Symposium for ICS & SCADA Cyber Security Research, Belfast, United Kingdom, 2016.

[66] Plattform Industrie 4.0, *Künstliche Intelligenz (KI) in Sicherheitsaspekten der Industrie 4.0*, German Federal Ministry for Economic Affairs and Energy (BMWi), 2019.

[67] A. Töpfer, *Erfolgreich Forschen: Ein Leitfaden für Bachelor-, Master-Studierende und Doktoranden*, Springer Gabler, 3. Auflage, 2012.

[68] B. Arbaugh, *Security: Technical, social, and legal Challenges*, Computer, Volume: 35, Issue: 2, 2002.

[69] National Security Agency and Cybersecurity and Infrastructure Security Agency, *Control System Defense: Know the Opponent*, Cybersecurity Advisory, 2022.

[70] S. Obermeier, *Cyber Security Research Challenges - An Industry Perspective*, 23rd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Torino, Italy, 2018.

[71] M. Maidl, D. Kröselberg, J. Christ, and K. Beckers, *A Comprehensive Framework for Security in Engineering Projects: Based on IEC 62443*, IEEE International Symposium on Software Reliability Engineering Workshops, 2018.

[72] National Institute of Standards and Technology, *Guide to Operational Technology (OT) Security*, Special Publication (SP) 800-82 Revision 3, 2022.

[73] International Society of Automation, *ISA-TR84.00.09-2023: Cyber Seurity Related to the Safety Lifecycle*, Technical Report, Rev 5.4, 3rd Edition, 2023.

[74] F. Patzer, *Automatisierte, minimalinvasive Sicherheitsanalyse und Vorfallreaktion für industrielle Systeme*, Karlsruher Schriften zur Anthropomatik, Band 56, 2021.

[75] M. Friesen, G. Karthikeyan, S. Heiss, L. Wisniewski, and H. Trsek, *A comparative Evaluation of Security Mechanisms in DDS, TLS, and DTLS*, 9. Jahreskolloquium Kommunikation in der Automation (KommA), Lemgo, Germany, 2018.

[76] S. Hollerer, W. Kastner, and T. Sauter, *Safety and Security: A Field of Tension in Industrial Practice*, e&i Elektrotechnik und Informationstechnik, 2021.

[77] M. Bouissou and L. Piètre-Cambacédès, *Modeling Safety and Security Interdependencies with BDMP (Boolean Logic Driven Markov Processes)*, IEEE International Conference on Systems, Man and Cybernetics, Istanbul, Turkey, 2010.

[78] P. Bhosale, W. Kastner, and T. Sauter, *Integrated Safety-Security Risk Assessment for Production Systems: A Use Case Using Bayesian Belief Networks*, IEEE 21st International Conference on Industrial Informatics (INDIN), Lemgo, Germany, 2023.

[79] S. Fluchs, *On Modelling for Security Engineering as a Submodel of the Digital Twin*, Fluchsfriction Blog, 2021.

[80] S. Paul and L. Rioux, *Over 20 Years of Research into Cybersecurity and Safety Engineering: A short Bibliography*, Safety and Security Engineering 5, 2015.

[81] P. Kleen, H. Flatt, and J. Jasperneite, *Erweiterung des „Secure Plug & Work“ für Safety-kritische Systeme*, VDI-Kongress Automation, Baden-Baden, Germany, 2017.

[82] D. Etz, T. Frühwirth, and W. Kastner, *Self-Configuring Safety Networks*, 9. Jahreskolloquium Kommunikation in der Automation (KommA), Lemgo, Germany, 2018.

[83] X. Lyu, Y. Ding, and S.-H. Yang, *Safety and Security Risk Assessment in Cyber-Physical Systems*, IET Cyber-Physical Systems: Theory Applications, 2019.

[84] E. Lisova, I. Šljivo, and A. Čaušević, *Safety and Security Co-Analyses: A Systematic Literature Review*, IEEE Systems Journal, Vol. 13, No. 3, 2019.

[85] C. Bieder and K. P. Gould, *The Coupling of Safety and Security: Exploring Interrelations in Theory and Practice*, Springer Briefs in Applied Sciences and Technology, Safety Management, 2020.

[86] S. Hollerer, W. Kastner, and T. Sauter, *Towards a Comprehensive Ontology Considering Safety, Security, and Operation Requirements in OT*, 28th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Sinaia, Romania, 2023.

[87] S. Kropatschek *et al.*, *Combining Models for Safety and Security Concerns in Automating Digital Production*, IEEE 21st International Conference on Industrial Informatics (INDIN), Lemgo, Germany, 2023.

[88] A. M. Hosseini, C. Fischer, M. Bhole, W. Kastner, T. Sauter, and S. Schlund, *A Safety and Security Requirements Management Methodology in Reconfigurable Collaborative Human-Robot Application*, 19th IEEE International Workshop on Factory Communication Systems (WFCS), Pavia, Italy, 2023.

[89] D. R. Coats, *Worldwide Threat Assessment of the US Intelligence Community*, Statement for the Record, 2018.

[90] World Economic Forum, *The Global Risks Report*, 15th Edition, 2020.

[91] IBM Security, *X-Force Threat Intelligence Index 2022*, IBM Corporation, Armonk, USA, 2022.

[92] U. P. D. Ani, H. He, and A. Tiwari, *Review of Cybersecurity Issues in Industrial Critical Infrastructure: Manufacturing in perspective*, Journal of Cyber Security Technology, Volume 1, Issue 1, 2016.

[93] A. Pattanayak and M. Kirkland, *Current Cyber Security Challenges in ICS*, IEEE International Conference on Industrial Internet (ICII), Seattle, USA, 2018.

[94] M. Ehrlich, A. Bröring, C. Diedrich, J. Jasperneite, W. Kastner, and H. Trsek, *Determining the Target Security Level for Automated Security Risk Assessments*, IEEE 21st International Conference on Industrial Informatics (INDIN), Lemgo, Germany, 2023.

[95] A. Di Pinto, Y. Dragoni, and A. Carcano, *TRITON: The First ICS Cyber Attack on Safety Instrument Systems*, Black Hat Rsearch Paper, USA, 2018.

[96] Dragos Inc., *Year in Review: ICS Vulnerabilities*, Annual Report, 2019.

[97] Dragos Inc., *Year in Review: The ICS Landscape and Threat Activity Groups*, Annual Report, 2019.

[98] CyberX, *Global IoT/ICS Risk Report: A Data-driven Analysis of Vulnerabilities in our Internet of Things (IoT) and Industrial Control Systems (ICS) Infrastructure*, 2020.

[99] Water Sector Coordinating Council, *Water and Wastewater Systems: Cybersecurity State of the Sector*, 2021.

[100] S. Hollerer, W. Kastner, and T. Sauter, *Safety and Security: A Field of Tension in Industrial Practice*, IEEE 21st International Conference on Industrial Informatics (INDIN), Lemgo, Germany, 2023.

[101] Dragos Inc., *Year in Review: Lessons Learned from the Front Lines of ICS Cybersecurity*, Annual Report, 2019.

[102] S. Zimmermann *et al.*, *Industrial Security in the Mechanical and Plant Engineering Industry: Results of the VDMA Study and Recommendations for Action*, VDMA Competence Center Industrial Security, 2019.

[103] TÜV Rheinland i-sec GmbH, *Industrial Security in 2019: A TÜV Rheinland Perspective*, Whitepaper, 2019.

[104] Ponemon Institute and Siemens, *Caught in the Crosshairs: Are Utilities Keeping Up with the Industrial Cyber Threat*, Smart Energy International, Issue 5, 2019.

[105] Dragos Inc., *Examining ICS/OT Exploits: Findings from more than a Decade of Data*, Whitepaper, 2021.

[106] Kaspersky, *IT-Budgets sinken, IT-Sicherheitsausgaben steigen*, Studie, 2017.

[107] IBM Security, *Cost of a Data Breach Report 2021*, IBM Corporation, Armonk, USA, 2021.

[108] KELA, *The Cybercrime Inferno: Annual Report on Ransomware, Extortion, and Network Access Sales*, KELA Cybercrime Prevention, 2022.

[109] B. Filkins, *SANS 2019 State of OT/ICS Cybersecurity Survey*, SANS Institute Information Security Reading Room, 2019.

[110] G. Wangen and E. Snekkenes, *A Taxonomy of Challenges in Information Security Risk Management*, Proceeding of Norwegian Information Security Conference (NISK), Stavanger, Norway, 2013.

[111] DIN/DKE, *DIN and DKE Roadmap: German Standardization Roadmap Industrie 4.0*, Version 4, 2020.

[112] M. Eckhart and A. Ekelhart, *Digital Twins for Cyber-Physical Systems Security: State of the Art and Outlook*, Springer Nature Switzerland AG, Security and Quality in Cyber-Physical Systems Engineering, 2019.

[113] M. Eckhart, B. Brenner, A. Ekelhart, and E. Weippl, *Quantitative Security Risk Assessment for Industrial Control Systems: Research Opportunities and Challenges*, Journal of Internet Services and Information Security (JISIS), Volume: 9, Issue: 3, 2019.

[114] M. Rocchetto, A. Ferrari, and V. Senni, *Challenges and Opportunities for Model-based Security Risk Assessment of Cyber-Physical Systems*, Resilience of Cyber-Physical Systems - Advanced Sciences and Technologies for Security Applications, Springer Nature, 2019.

[115] J. Wetzels, D. dos Santos, and M. Ghafari, *Insecure by Design in the Backbone of Critical Infrastructure*, IEEE/ACM Workshop on the Internet of Safe Things, San Antonio, USA, 2023.

[116] A. Kuppa, L. Aouad, and N.-A. Le-Khac, *Linking CVE's to MITRE ATT&CK Techniques*, 16th International Conference on Availability, Reliability and Security (ARES), Vienna, Austria, 2021.

[117] E. Bergström, M. Lundgren, and Å. Ericson, *Revisiting Information Security Risk Management Challenges: A practice perspective*, Information & Computer Security, 2019.

[118] Y. Barlette and V. V. Fomin, *Exploring the Suitability of IS Security Management Standards for SMEs*, 41st Hawaii International Conference on System Sciences (HICSS), Waikoloa, USA, 2008.

[119] C. G. Manso, E. Rekleitis, F. Papazafeiropoulos, and V. Maritsas, *Information Security and Privacy Standards for SMEs: Recommendations to improve the Adoption of Information Security and Privacy Standards in Small and Medium Enterprises*, European Union Agency for Network and Information Security (ENISA), 2015.

[120] C. Tebbe, M. Glawe, K.-H. Niemann, and A. Fay, *Informationsbedarf für automatische IT-Sicherheitsanalysen automatisierungstechnischer Anlagen*, at – Automatisierungstechnik, Volume: 65, Issue: 1, 2017.

[121] M. Ehrlich, H. Trsek, L. Wisniewski, and J. Jasperneite, *Survey of Security Standards for an automated Industrie 4.0 compatible Manufacturing*, 45th Annual Conference of the IEEE Industrial Electronics Society (IECON), Lisbon, Portugal, 2019.

[122] H. Trsek, D. Mahrenholz, S. Schemmer, and R. Schumann, *Industrial Security 4.0 - Zukünftige Herausforderungen und Lösungen zur Absicherung von Cyber-physischen Produktionssystemen*, Markt & Technik Industrie 4.0 & Industrial Internet Summit und Security Symposium, Munich, Germany, 2015.

[123] C. Montes Portela, M. Hoeve, F. H. Tan, and H. Slootweg, *Implementing an ISA/IEC 62243 and ISO/IEC 27001 OT Cyber Security Management System at Dutch DSO Enexis*, 25th International Conference on Electricity Distribution (CIRED), Madrid, Spain, 2019.

[124] T. Sommestad, G. N. Ericsson, and J. Nordlander, *SCADA System Cyber Security: A Comparison of Standards*, IEEE PES General Meeting, 2010.

[125] I. Rolle, *Funktionale Sicherheit: Modelle der funktionalen Sicherheit*, Elektronik Magazin 8/2013, 2013.

[126] K. Beckers, *Pattern and Security Requirements: Engineering-Based Establishment of Security Standards*, Springer International Publishing Switzerland, 2015.

[127] C. Siegwart, H. Adamczyk, and G. Frey, *Industrial Security: I4.0 Analysis of the IEC 62443*, VDI-Kongress Automation, Baden-Baden, Germany, 2018.

[128] S. Fluchs and H. Rudolph, *Making OT Security Engineering an Engineering Discipline: A method-agnostic Thought Model and a Data Model*, atp, Transforming Automation, 2019.

[129] D. Bernoulli, *Exposition of a New Theory on the Measurement of Risk*, Econometrica, Vol. 22, No. 1., 1954.

[130] S. Kaplan and B. J. Garrick, *On The Quantitative Definition of Risk*, Risk Analysis, Vol. 1, No. 1, 1981.

[131] Committee to Review the Department of Homeland, *Review of th Department of Homeland Security's Approach to Risk Analysis*, National Rsearch Council, 2010.

[132] European Cyber Security Organisation, *State of the Art Syllabus: Overview of existing Cybersecurity Standards and Certification Schemes*, WG1 – Standardisation, certification, labelling and supply chain management, v2, 2017.

[133] European Cyber Security Organisation, *European Cyber Security Certification: A Meta-Scheme Approach*, WG1 – Standardisation, certification, labelling and supply chain management, v1, 2017.

[134] European Cyber Security Organisation, *The ECSO Cybersecurity Market Radar*, Whitepaper, 2019.

[135] M. Gergeleit, H. Trsek, T. Eisert, and M. Ehrlich, *Modelling Security Requirements and Controls for an Automated Deployment of Industrial IT Systems*, 9. Jahreskolloquium Kommunikation in der Automation (KommA), Lemgo, Germany, 2018.

[136] B. Kime, *Now Tech: Industrial Control Systems (ICS) Security Solutions*, Forrester Whitepaper, 2021.

[137] R. Leszczyna, *Review of Cybersecurity Assessment Methods: Applicability Perspective*, Computers & Security, Volume 108, 2021.

[138] T. Gamer, B. Kloepper, and M. Hoernicke, *The Way Toward Autonomy in Industry: Taxonomy, Process Framework, Enablers, and Implications*, 45th Annual Conference of the IEEE Industrial Electronics Society (IECON), Lisbon, Portugal, 2019.

[139] National Institute of Standards and Technology, *Security and Privacy Controls for Information Systems and Organizations*, Draft NIST Special Publication 800-53, Revision 5, 2017.

[140] L. Lemaire, J. Vossaert, B. de Decker, and V. Naessens, *An Assessment of Security Analysis Tools for Cyber-Physical Systems*, International Workshop on Risk Assessment and Risk-Driven Testing, 2017.

[141] C. D. Jaeger, N. S. Roehrig, and T. Torres, *Development of an Automated Security Risk Assessment Methodology Tool for Critical Infrastructures*, Sandia National Laboratories Report, 2008.

[142] R. Schlegel, S. Obermeier, and J. Schneider, *Structured System Threat Modeling and Mitigation Analysis for Industrial Automation Systems*, IEEE 13th International Conference on Industrial Informatics (INDIN), Cambridge, United Kingdom, 2015.

[143] P. Smith, *Smart Grid Protection Against Cyber Attacks (SPARKS): Threat and Risk Assessment Methodology*, Project Deliverable 2.2, 2015.

[144] J. Eichler and D. Angermeier, *Modular Risk Assessment for the Development of secure Automotive Systems*, 31. VDI/VW-Gemeinschaftstagung Automotive Security, Wolfsburg, Germany, 2015.

[145] M. Kern, T. Glock, V. P. Betancourt, B. Liu, J. Becker, and E. Sax, *Model-based Support to increase Traceability in the Context of Security-by-Design for I4.0 Systems*, VDI-Kongress Automation, Baden-Baden, Germany, 2020.

[146] M. Eckhart, A. Ekelhart, and E. Weippl, *Automated Security Risk Identification Using AutomationML-based Engineering Data*, IEEE Transactions on Dependable and Secure Computing, 2020.

[147] A. Couce-Vieira, S. H. Houmb, and D. Ríos-Insua, *CSIRA: A Method for Analysing the Risk of Cybersecurity Incidents*, 4th International Workshop GraMSec, Santa Barbara, USA, 2018.

[148] B. Leander, A. Čaušević, and H. Hansson, *Cybersecurity Challenges in Large Industrial IoT Systems*, 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain, 2019.

[149] B. Fabian, S. Gürses, M. Heisel, T. Santen, and H. Schmidt, *A Comparison of Security Requirements Engineering Methods*, Special Issue - Security Requirements Engineering, Springer-Verlag London Limited, 2010.

[150] D. Gritzalis, G. Iseppi, A. Mylonas, and V. Stavrou, *Exiting the Risk Assessment Maze: A Meta-Survey*, ACM Computing Surveys 51, 1, Article 11, 2018.

[151] M. Kern *et al.*, *An Architecture-based Modeling Approach Using Data Flows for Zone Concepts in Industry 4.0*, IEEE International Symposium on Systems Engineering (ISSE), Vienna, Austria, 2020.

[152] J. Bau and J. C. Mitchell, *The Science of Security: Security Modeling and Analysis*, IEEE Security & Privacy, Volume: 9, Issue: 3, 2011.

[153] B. Kordy, L. Piètre-Cambacédès, and P. Schweitzer, *DAG-Based Attack and Defense Modeling: Don't Miss the Forest for the Attack Trees*, Computer Science Review, Volumes 13-14, 2014.

[154] A. Ramos, M. Lazar, R. H. Filho, and Rodrigues, Joel J. P. C., *Model-Based Quantitative Network Security Metrics: A Survey*, IEEE Communications Surveys & Tutorials, Volume: 19, Issue: 4, 2017.

[155] H. Orojloo and M. A. Azgomi, *A Method for Modeling and Evaluation of the Security of Cyber-Physical Systems*, 11th International ISC Conference on Information Security and Cryptology, Tehran, Iran, 2014.

[156] Y. Peng, K. Huang, W. Tu, and C. Zhou, *A Model-Data Integrated Cyber Security Risk Assessment Method for Industrial Control Systems*, IEEE 7th Data Driven Control and Learning Systems Conference (DDCLS), Enshi, China, 2018.

[157] P. Bhosale, W. Kastner, and T. Sauter, *AutomationML Use for Safety and Security Risk Assessment in Industrial Control Systems*, 28th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Sinaia, Romania, 2023.

[158] C. Hildebrandt *et al.*, *Semantic Modeling for Collaboration and Cooperation of Systems in the Production Domain*, 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol, Cyprus, 2017.

[159] A. M. Hosseini, T. Sauter, and W. Kastner, *Safety and Security Requirements in AAS Integration: Use Case Demonstration*, 19th IEEE International Workshop on Factory Communication Systems (WFCS), Pavia, Italy, 2023.

[160] S. Fenz and A. Ekelhart, *Formalizing Information Security Knowledge*, Proceedings of the 4th International Symposium on Information, Computer and Communications Security (ASIACCS), Sydney, Australia, 2009.

[161] J. D. Howard and T. A. Longstaff, *A Common Language for Computer Security Incidents*, Sandia National Laboratories Report, 1998.

[162] D. Angermeier, K. Beilke, G. Hansch, and J. Eichler, *Modeling Security Risk Assessments*, Embedding Security in Cars (escar Europe), Stuttgart, Germany, 2019.

[163] A. Bunte, O. Niggemann, and B. Stein, *Integrating OWL Ontologies for Smart Services into AutomationML and OPC UA*, 23rd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Torino, Italy, 2018.

[164] Y. Merah and T. Kenaza, *Ontology-based Cyber Risk Monitoring Using Cyber Threat Intelligence*, 16th International Conference on Availability, Reliability and Security (ARES), Vienna, Austria, 2021.

[165] I. Grangel-González, L. Halilaj, G. Coskun, S. Auer, D. Collarana, and M. Hoffmeister, *Towards a Semantic Administrative Shell for Industry 4.0 Components*, Tenth International Conference on Semantic Computing (SC), 2016.

[166] E. Kiesling, A. Ekelhart, K. Kurniawan, and F. Ekaputra, *The SEPSES Knowledge Graph: An Integrated Resource for Cybersecurity*, International Semantic Web Conference (ISWC), Auckland, New Zealand, 2019.

[167] C. Grigoriadis, A. M. Berzovitis, I. Stellios, and P. Kotzanikolaou, *A Cybersecurity Ontology to Support Risk Information Gatherin in Cyber-Physical Systems*, Springer Nature Switzerland AG, ESORICS International Workshops, 2022.

[168] G. Hansch, *Automating Security Risk and Requirements Management for Cyber-Physical Systems*, Dissertation, Promotionsprogramm Computer Science (PCS) der Georg-August University School of Science (GAUSS), 2020.

[169] J. Wolf, *Security of Systems: Modelling and Analysis Methodology*, Dissertation, Fakultät für Mathematik/Informatik und Maschinenbau der Technischen Universität Clausthal, Clausthal-Zellerfeld, 2022.

[170] B. A. Mozzaquatro, R. Melo, C. Agostinho, and R. Jardim-Goncalves, *An Ontology-Based Security Framework for Decision-Making in Industrial Systems*, 4th International Conference on Model-Driven Engineering and Software Development (MODELSWARD), Rome, Italy, 2016.

[171] M. Almorsy and J. Grundy, *SecDSVL: A Domain-Specific Visual Language To Support Enterprise Security Modelling*, 23rd Australian Software Engineering Conference, Milsons Point, Australia, 2014.

[172] S. Hollerer, T. Sauter, and W. Kastner, *Risk Assessments Considering Safety, Security, and Their Interdependencies in OT Environments*, The 17th International Conference on Availability, Reliability and Security (ARES), Vienna, Austria, 2022.

[173] M. J. M. Chowdhury, *Security Risk Modelling Using SecureUML*, 16th International Conference on Computer and Information Technology (ICCIT), Khulna, Bangladesh, 2014.

[174] D. A. Robles-Ramirez, P. J. Escamilla-Ambrosio, and T. Tryfonas, *IoTsec: UML Extension for Internet of Things Systems Security Modelling*, International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE), Cuernavaca, Mexico, 2017.

[175] M. Eckhart, A. Ekelhart, S. Biffl, A. Lüder, and E. Weippl, *QualSec: An Automated Quality-Driven Approach for Security Risk Identification in Cyber-Physical Production Systems*, IEEE Transactions on Industrial Informatics, Volume: 19, Issue: 4, 2022.

[176] D. Mažeika and R. Butleris, *Integrating Security Requirements Engineering into MBSE: Profile and Guidelines*, Security and Communication Networks, 2020.

[177] M. Bhole, W. Kastner, and T. Sauter, *A Model Based Framework for Testing Safety and Security in Operational Technology Environments*, 27th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Stuttgart, Germany, 2022.

[178] C. Beierle and G. Kern-Isberner, *Methoden wissensbasierter Systeme: Grundlagen, Algorithmen, Anwendungen*, 6. Edition, Computational Intelligence, Springer Vieweg, 2019.

[179] M. Ehrlich, G. Lukas, H. Trsek, J. Jasperneite, W. Kastner, and C. Diedrich, *Requirements Analysis for the Evaluation of Automated Security Risk Assessments*, 20th IEEE International Workshop on Factory Communication Systems (WFCS), Toulouse, France, 2024.

[180] C. Tebbe, K.-H. Niemann, and A. Fay, *Ontology and Life Cycle of Knowledge for ICS Security Assessments*, 4th International Symposium for ICS & SCADA Cyber Security Research, Belfast, United Kingdom, 2016.

[181] M. Ehrlich, A. Biendarra, H. Trsek, E. Wojtkowiak, and J. Jasperneite, *Passive Flow Monitoring of Hybrid Network Connections regarding Quality of Service Parameters for the Industrial Automation*, 8. Jahreskolloquium Kommunikation in der Automation (KommA), Magdeburg, Germany, 2017.

[182] S. Fluchs and H. Rudolph, *Making OT Security Engineering deserve its Name*, Control Global Magazine, 2019.

[183] E. Tastan, S. Fluchs, and R. Drath, *Warum wir ein Security-Engineering-Informationsmodell brauchen: Motivation, Anwendungsfälle und Konzept für ein neues Domänenmodell für Security-Engineering*, 18. AALE-Konferenz, Pforzheim, Germany, 2022.

[184] S. Fluchs, E. Tastan, T. Trumpf, A. Horch, R. Drath, and A. Fay, *Traceable Security-by-Design Decisions for Cyber-Physical Systems (CPSs) by Means of Function-Based Diagrams and Security Libraries*, Sensors 2023, Special Issue, IoT-Based Cyber-Physical System: Challenges and Future Direction, 2023.

[185] B. Brenner *et al.*, *Better safe than sorry: Risk Management based on a safety-augmented Network Intrusion Detection System*, IEEE Open Journal of the Industrial Electronics Society, Volume: 4, 2023.

[186] A. Ekelhart, S. Fenz, and T. Neubauer, *AURUM: A Framework for Information Security Risk Management*, 42nd Hawaii International Conference on System Sciences, Waikoloa, USA, 2009.

[187] Z. Anwar and R. Campbell, *Automated Assessment of Compliance with Security Best Practices*, International Federation for Information Processing (IFIP), Volume 290, 2008.

[188] T. He and Z. Li, *A Model and Method of Information Security Risk Assessment based on MITRE ATT&CK*, 2nd International Conference on Electronics, Communications and Information Technology (CECIT), Sanya, China, 2021.

[189] M. Chorąś, A. Flizikowski, and R. Kozik, *Decision Aid Tool and Ontology-Based Reasoning for Critical Infrastructure Vulnerabilities and Threats Analysis*, Critical Information Infrastrucutres Security (CRITIS), 2010.

[190] Y. Hua and B. Hein, *Interpreting OWL Complex Classes in AutomationML based on Bidirectional Translation*, 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain, 2019.

[191] O. Cardin, *Classification of Cyber-Physical Production Systems Applications: Proposition of an Analysis Framework*, Computers in Industry, Volume 104, 2019.

[192] S. Hollerer, W. Kastner, and T. Sauter, *Towards a Threat Modeling Approach Addressing Security and Safety in OT Environments*, 17th IEEE International Workshop on Factory Communication Systems (WFCS), Linz, Austria, 2021.

[193] X. Ou, *A Logic-Programming Approach to Network Security Analysis*, Dissertation, Department of Computer Science, Faculty of Princeton University, 2005.

[194] S. Fluchs *et al.*, *Integrating Security Decisions "by design" into the Engineering of Process Plants: Introducing the "Automation Security by Design Decisions" Concept*, VDI-Kongress Automation, Baden-Baden, Germany, 2022.

[195] J. D. Gilsinn and R. Schierholz, *Security Assurance Levels: A Vector Approach to Describing Security Requirements*, NIST Publication, 2010.

[196] S. Fenz and A. Ekelhart, *Verification, Validation, and Evaluation in Information Security Risk Management*, IEEE Security & Privacy, Volume: 9, Issue: 2, 2011.

[197] P. Wagner, G. Hansch, C. Konrad, K.-H. John, J. Bauer, and J. Franke, *Applicability of Security Standards for Operational Technology by SMEs and Large Enterprises*, 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria, 2020.

[198] Object Management Group, *OMG Unified Modeling Language*, Version 2.5.1, 2017.

[199] P. Bhosale, W. Kastner, and T. Sauter, *Automating Safety and Security Risk Assessment in Industrial Control Systems: Challenges and Constraints*, 27th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Stuttgart, Germany, 2022.

[200] Plattform Industrie 4.0, *Details of the Asset Administration Shell Part 1: The Exchange of Information between Partners in the Value Chain of Industrie 4.0*, German Federal Ministry for Economic Affairs and Energy (BMWi), Version 3.0RC02, 2020.

[201] D. Snowden, *The ASHEN Model: An Enabler of Action*, The Cynefin Centre, 2005.

[202] S. Choi, J.-H. Yun, and S.-K. Kim, *A Comparison of ICS Datasets for Security Research Based on Attack Paths*, International Conference on Critical Information Infrastructure Security, Kaunas, Lithuania, 2018.

[203] J. DeTreville, *Binder, a Logic-based Security Language*, Proceedings 2002 IEEE Symposium on Security and Privacy, Berkeley, USA, 2002.

[204] S. Rao, *A Foundation for System Safety Using Predicate Logic*, 3rd Annual IEEE Systems Conference, Vancouver, Canada, 2009.

[205] M. Shafto *et al.*, *Modeling, Simulation, Information Technology & Processing Roadmap*, Draft, Technology Area 11, National Aeronautics and Space Administration (NASA), 2010.

[206] R. Drath *et al.*, *Interoperabilität mit der Verwaltungsschale, OPC UA und AutomationML: Zielbild und Handlungsempfehlungen für industrielle Interoperabilität*, Diskussionspapier, Industrial Digital Twin Association, 2023.

[207] A. Bröring, M. Ehrlich, H. Trsek, and L. Wisniewski, *Secure Usage of Asset Administration Shells: An Overview and Analysis of Best Practises*, 12. Jahreskolloquium Kommunikation in der Automation (KommA), Magdeburg, Germany, 2021.

[208] F. Patzer, F. Volz, T. Usländer, I. Blöcher, and J. Beyerer, *The Industrie 4.0 Asset Administration Shell as Information Source for Security Analysis*, 24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Zaragoza, Spain, 2019.

[209] C. Diedrich *et al.*, *Sprache für I4.0-Komponenten: Semantik der Interaktionen von I4.0 Komponenten*, VDI-Kongress Automation, Baden-Baden, Germany, 2018.

[210] A. Bröring, M. Ehrlich, L. Wisniewski, H. Trsek, and S. Heiss, *An Asset Administration Shell Version Control to Enforce Integrity Protection*, 13. Jahreskolloquium Kommunikation in der Automation (KommA), Lemgo, Germany, 2022.

[211] M. Jacoby *et al.*, *Open-Source Implementations of the Reactive Asset Administration Shell: A Survey*, Sensors, 23, 5229, 2023.

[212] L. Stojanovic *et al.*, *Methodology and Tools for Digital Twin Management: The FA^{3}ST Approach*, IoT, 2, 2021.

[213] M. Eckhart *et al.*, *Security-Enhancing Digital Twins: Characteristics, Indicators, and Future Perspectives*, IEEE Security & Privacy, Volume: 21, Issue: 6, 2023.

[214] A. Bröring, M. Ehrlich, L. Wisniewski, H. Trsek, and S. Heiss, *Towards an Asset Administration Shell Integrity Verification Scheme*, 27th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Stuttgart, Germany, 2022.

[215] Verein Deutscher Ingenieure, *Modulare Anlagen: Paradigmenwechsel im Anlagenbau: Zusammenspiel von Prozesstechnik und Automatisierungstechnik*, VDI-Handlungsempfehlung, 2022.

[216] S. Grüner *et al.*, *Integration of Module Type Package and Industry 4.0 AssetAdministration Shell*, VDI-Kongress Automation, Baden-Baden, Germany, 2021.

[217] C. Wagner *et al.*, *The Role of the Industry 4.0 Asset Administration Shell and the Digital Twin during the Life Cycle of a Plant*, 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Limassol, Cyprus, 2017.

[218] A. M. Hosseini, T. Sauter, and W. Kastner, *Towards Adding Safety and Security Properties to the Industry 4.0 Asset Administration Shell*, 17th IEEE International Workshop on Factory Communication Systems (WFCS), Linz, Austria, 2021.

[219] A. M. Hosseini, T. Sauter, and W. Kastner, *A Safety and Security Reference Architecture for Asset Administration Shell Design*, 18th IEEE International Workshop on Factory Communication Systems (WFCS), Pavia, Italy, 2022.

[220] X. Ye, W. S. Song, S. H. Hong, Y. C. Kim, and N. H. Yoo, *Toward Data Interoperability of Enterprise and Control Applications via the Industry 4.0 Asset Administration Shell*, IEEE Access, 2022.

[221] P. Bhosale, W. Kastner, and T. Sauter, *Integrated Safety-Security Risk Assessment for Industrial Control System: An Ontology-based Approach*, 28th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Sinaia, Romania, 2023.

[222] Industrial Digital Twin Association, *Specification of the Asset Administration Shell: Part 1: Metamodel*, Specification, IDTA Number: 01001-3-0, Industrial Digital Twin Association, 2023.

[223] Plattform Industrie 4.0, *Details of the Asset Administration Shell Part 2: Interoperability at Runtime - Exchanging Information via Application Programming Interfaces*, German Federal Ministry for Economic Affairs and Energy (BMWi), Version 1.0RC02, 2021.

[224] R. Pribiš, L. Beňo, and P. Drahoš, *Asset Administration Shell Design Methodology Using Embedded OPC Unified Architecture Server*, Electronics, 10, 2520, 2021.

[225] T. Miny, M. Thies, S. Heppner, I. Garmaev, L. Möller, and T. Kleinert, *Realisierung und Evaluation des Verwaltungsschalen-Metamodells: Python-SDK PyI40AAS und Lösungen für die praktische Umsetzung*, atp magazin 63(9), 2022.

[226] D. Berardi, F. Callegati, A. Giovine, A. Melis, M. Prandini, and L. Rinieri, *When Operation Technology Meets Information Technology: Challenges and Opportunities*, Future Internet, 15, 95, 2023.

[227] M. Ehrlich, A. Bröring, H. Trsek, J. Jasperneite, and C. Diedrich, *Evaluation Concept for Prototypical Implementation towards Automated Security Risk Assessments*, 28th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Sinaia, Romania, 2023.

[228] R. Sherwood *et al.*, *Can the Production Network be the Testbed?* 9th USENIX conference on Operating systems design and implementation (OSDI), Vancouver, Canada, 2010.

[229] M. Ehrlich, M. Gergeleit, H. Trsek, and G. Lukas, *Towards Automated Security Evaluation within the Industrial Reference Architecture*, 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Vienna, Austria, 2020.

[230] O. Andreeva *et al.*, *Industrial Control Systems and their online Availability*, Kaspersky Lab Report, 2016.

[231] N. Naik, P. Jenkins, P. Grace, and J. Song, *Comparing Attack Models for IT Systems: Lockheed Martin's Cyber Kill Chain, MITRE ATT&CK Framework and Diamond Model*, 8th IEEE International Symposium on Systems Engineering (ISSE), Vienna, Austria, 2022.

[232] O. Alexander, M. Belisle, and J. Steele, *MITRE ATT&CK for Industrial Control Systems: Design and Philosophy*, MITRE Corporation, Originally Published March 2020, 2020.

[233] B. E. Strom, A. Applebaum, D. P. Miller, K. C. Nickels, A. G. Pennington, and C. B. Thomas, *MITRE ATT&CK: Design and Philosophy*, MITRE Corporation, Originally Published July 2018, Revised March 2020, 2020.

[234] M. Bristow, *A SANS 2021 Survey: OT/ICS Cybersecurity*, SANS Institute Analyst Program, 2021.

[235] T. Casey, *Threat Agent Library Helps Identify Information Security Risks*, Intel Information Technology White Paper, 2007.

[236] T. Casey, P. Koeberl, and C. Vishik, *Threat Agents: A Necessary Component of Threat Analysis*, Proceedings of the Sixth Annual Workshop on Cyber Security and Information Intelligence Research, 2010.

[237] J. Bugeja, A. Jacobsson, and P. Davidsson, *An Analysis of Malicious Threat Agents for the Smart Connected Home*, IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), 2017.

[238] M. Ehrlich *et al.*, *Software-Defined Networking as an Enabler for Future Industrial Network Management*, 23rd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), Torino, Italy, 2018.

[239] M. Ehrlich, D. Lang, M. Mytych, T. Malischewski, and J. Jasperneite, *Preparing Smart Sensors for Industrie 4.0: Requirements, Potentials, and Solutions*, VDI-Kongress Automation, Baden-Baden, Germany, 2018.