Model-based Trustworthiness Evaluation of Autonomous Cyber-Physical Production Systems: A Systematic Mapping Study

MARYAM ZAHID, ALESSIO BUCAIONI, and FRANCESCO FLAMMINI, Mälardalen University, Eskilstuna, Sweden

The fourth industrial revolution, i.e., Industry 4.0, is associated with Cyber-Physical Systems (CPS), which are entities integrating hardware (e.g., smart sensors and actuators connected through the Industrial Internet of Things) together with control and analytics software used to drive and support decisions at several levels. The latest developments in Artificial Intelligence (AI) and Machine Learning (ML) have enabled increased autonomy and closer human-robot cooperation in the production and manufacturing industry, thus leading to Autonomous Cyber-Physical Production Systems (ACPPS) and paving the way to the fifth industrial revolution (i.e., Industry 5.0). ACPPS are increasingly critical due to the possible consequences of their malfunctions on human co-workers, and therefore, evaluating their trustworthiness is essential. This article reviews research trends, relevant attributes, modeling languages, and tools related to the model-based trustworthiness evaluation of ACPPS. As in many other engineering disciplines and domains, model-based approaches, including stochastic and formal analysis tools, are essential to master the increasing complexity and criticality of ACPPS and to prove relevant attributes such as system safety in the presence of intelligent behaviors and uncertainties.

CCS Concepts: • **Software and its engineering** → *Software verification and validation*; **Extra-functional properties**; • **Computer systems organization** → **Embedded and cyber-physical systems**; • **General and reference** → Surveys and overviews;

Additional Key Words and Phrases: Autonomous cyber-physical production systems, cyber-physical manufacturing systems, industry 4.0, industry 5.0, automation, trustworthiness, models, mapping study

ACM Reference Format:
Maryam Zahid, Alessio Bucaioni, and Francesco Flammini. 2024. Model-based Trustworthiness Evaluation of Autonomous Cyber-Physical Production Systems: A Systematic Mapping Study. *ACM Comput. Surv.* 56, 6, Article 157 (February 2024), 28 pages. https://doi.org/10.1145/3640314

1 INTRODUCTION

The concept of **Cyber-Physical Systems** (**CPS**) emerged with the 4th industrial revolution (see Figure 1) [58]. CPS are intelligent and networked systems, the development of which is characterized by close cooperation of its mechanical, electrical, and software components [16]. Nowadays,

M. Zahid, A. Bucaioni, and F. Flammini contributed equally to this research.
Authors' address: M. Zahid and A. Bucaioni, Mälardalen University, Hamngatan 15, Eskilstuna, 632 17, Sweden; e-mails: maryam.zahid@mdu.se, alessio.bucaioni@mdu.se; F. Flammini, Mälardalen University, Hamngatan 15, Eskilstuna, 632 17, Sweden, Department of Innovative Technologies, University of Applied Sciences and Arts of Sothern Switzerland, Lugano, CH-6900, Switzerland; e-mail: francesco.flammini@mdu.se.

ACM 0360-0300/2024/02-ART157
https://doi.org/10.1145/3640314

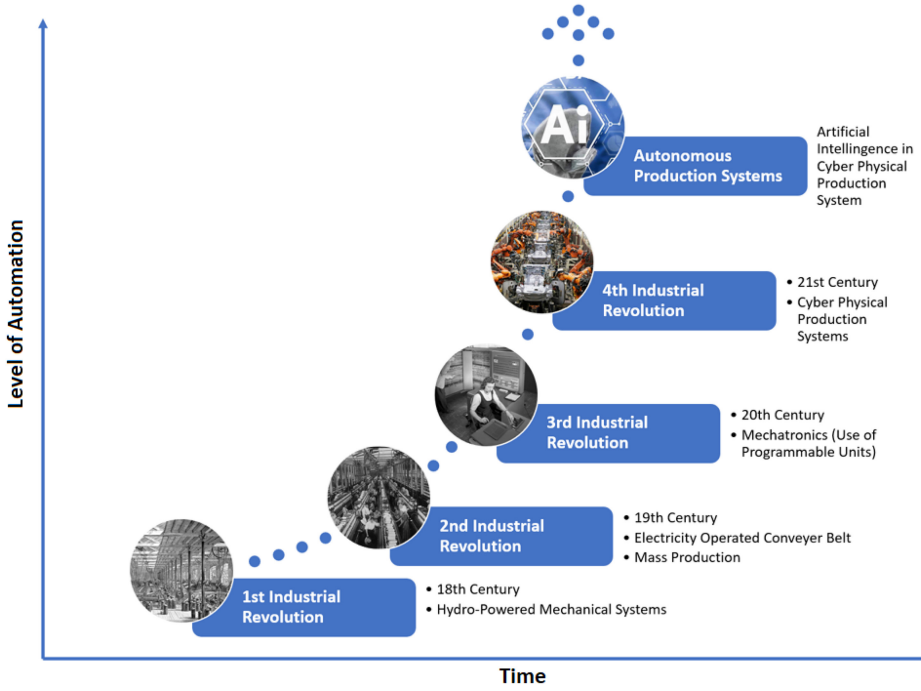

Fig. 1. Industrial revolutions over time leading to current ACPPS.

many CPS applications exist in diverse domains, such as smart grids, **industrial control systems** (**ICS**), autonomous vehicles, smart medical devices, and smart wearable devices.

In the production industry, those systems support process and/or manufacturing operations and can be hence defined as **cyber-physical production systems** (**CPPS**). CPPS is a specialized subset of CPS where the integrated features of CPS in the production process allow for simplifying the handling of many design and development challenges due to their ability to connect the cyber and physical components. CPS here integrates the physical process of the production system with the digital system, thus enhancing the manufacturing efficiency, flexibility, and quality [66]. Improving the interactions between system components, including machines, sensors, and information systems, enables more autonomy, agility, and reliability. Therefore, the traditional manufacturing process was transformed into a flexible process, which is more efficient and easily adaptable to the market trends [86].

According to Figure 1, the revolution in industrial processes did not just come to a halt with the introduction of CPS. A further revolution took place in the form of bringing about the concept of Industry 5.0 [24, 29] through the integration of **artificial intelligence** (**AI**) in CPPS, e.g., in the form of Cobots (Collaborative Robots) and other AI technologies being introduced [63], transforming traditional CPPS into smart and **autonomous CPPS** (**ACPPS**). Cobots are collaborative AI agents introduced in product manufacturing industries that perform tedious tasks in possibly dull and hazardous environments [72]. Using cobots can maximize production while addressing the varying demands of the targeted market, thus leveraging automation in CPPS [24, 74].

Some of the key features of ACPPS include the use of the digital twin [38, 39] or digital shadow to construct a virtual representation of the physical system enabling simulations, testings, and verification of the system's behavior and overall performance under varying conditions [77]. Its networked nature enables instant access to a large amount of data coming in from a network of

smart machines, sensors, actuators and controllers communicating and cooperating among each other and with humans [76]. ACPPS uses this information to generate information or knowledge about its environment and the actors in it. ACPPS uses the collected data and machine learning algorithms to generate and continuously update the predictive models, conduct root cause analysis, and parameter optimization for decision support or autonomous control functionalities [8].

Due to its vicinity with humans and no cages to isolate them, in addition to ethical issues, cobots pose challenges regarding predictability and consequences of any harmful misbehavior possibly causing injuries or even deaths to their human co-workers.

The production industry requires high availability, sufficient performance, and continuous data exchange to enable flexibility in possibly safety-critical ACPPS. That requires ACPPS to be considered trustworthy, such that there exists a minimum probability of disturbance or error during the execution of the process [16]. Depending on reference taxonomies, trustworthiness may be associated not only with aspects and attributes such as robustness, availability, reliability, safety, and security, which are included in the integrated concept of resilience [15] (i.e., dependability in the presence of changes and uncertainties), but it can also include ethical and legal aspects [30, 75], which we are not addressing in this study. Regarding legal aspects, evaluation of processes using model-based techniques often seems to be needed to fulfill certification requirements of international standards [53, 55].

It is worth mentioning that within ACPPS, lack of system security, i.e., adequately implemented and demonstrated protection against intentional attacks, can also lead to safety incidents [52], possibly resulting in catastrophic consequences on the system, its environment, and people involved [35]. Susceptibility to cyber-attacks can also result in a modification to the production process, leading to the manufacturing of incorrect parts of the final product, thus making it hazardous to its users and its environment [86]. Therefore, ensuring and proving trustworthiness in the system is of utmost importance.

Implementing proper fault-tolerance mechanisms is not enough to ensure trust in the system; in other words, common industry verification and validation processes also require engineers to evaluate the relevant aspects and attributes of trustworthiness formally. Since ACPPS is a relatively new domain, many code-based and model-based evaluation techniques have been proposed over the years but are mainly for CPPS and not for ACPPS. Moreover, it is believed that the code-based evaluation techniques do fulfill the purpose of evaluating the systems effectively but are costly in terms of development and maintenance costs. In such a context, model-based evaluation techniques allow early detection of faults in the system and effectively support certification activities against international security and safety standards [73].

Ensuring trustworthiness in ACPPS has become crucial. On the other hand, many challenges related to its evaluation have also been gaining interest in the research community and among practitioners. This highlighted the need for a systematic study to provide an in-depth insight into the current trends adopted to evaluate trustworthiness in ACPPS through model-based engineering approaches.

To the best of our knowledge, several mapping studies, literature reviews, and surveys exist addressing the evaluation of CPPS trustworthiness attributes. However, they are neither focused on ACPPS nor address the possible use of model-based evaluation techniques. Since ACPPS is a relatively new concept and is gaining popularity amongst various types of manufacturing industries, there is a need to investigate the use of evaluation mechanisms and the extent to which these techniques help evaluate the system.

Given that motivation, this study presents the planning, execution, and results of a systematic mapping study, providing a structured and comprehensive overview of trustworthiness evaluation in ACPPS by using model-based techniques. Starting from an initial set of 959 peer-review

publications, we came across 44 potential *primary* studies, which were analyzed following a meticulous data extraction, analysis, and synthesis process. Furthermore, this study also presents an orthogonal analysis of the techniques and their supporting tools used against the trustworthiness-related attributes being evaluated.

Below is a summary of our main findings from this study:

- Although the term "Cyber-Physical System" was coined in the year 2006, its introduction into the manufacturing industry together with AI and data analytics approaches made its first appearance in peer-reviewed publications much later: according to our study, the first instance of ACPPS being specifically evaluated for trustworthiness by using model-based techniques is dated 2016. The domain reached its popularity in the year 2020, with most articles published being conference papers, demonstrating the domain's novelty that is still mostly unexplored, especially considering aspects such as formally modeling intelligent behaviors to demonstrate compliance against relevant safety and security standards.
- Trustworthiness may include several aspects, but not all of them have been equally interesting to the research community. Security and safety attributes have been the most evaluated aspects in ACPPS, whereas performability (i.e., performance in the presence of dependability threats) and the more general, integrated concepts and holistic resilience metrics were addressed only in a limited number of studies.
- Even though all the articles studied for this mapping study proposed a model-based technique for evaluating trustworthiness, only 39 articles actually presented the working of their proposed technique using a case study related to the domain of ACPPS. Some used pre-existing tools, while others developed their own tools for evaluation purposes. Among these tools, Petri Net IDE (Integrated Development Environments) was the most reported tool, followed by MATLAB, which supports diverse modeling formalisms. Some of the tools used to model specific ACPPS components to be evaluated, while others used were specifically designed for evaluation purposes.
- Models instantiated from the used modeling languages were either used to present the behavior and structural aspects of the entire system or its component to be evaluated. Amid all this, some used modeling languages that provided means for evaluating the given system or its components. Some modeling approaches are suggested by the reference cyber-security standards specific to the production systems (such as ISO standards or IEC standards, e.g., ISA/IEC 62443 a standard for security in **Industrial Automation Control Systems** (**IACS**) [53]), while others have been proposed outside of those recommended by these standards as evaluation techniques. Various types of trees (such as Decision Trees, Attack Trees, and Fault Trees), Petri Nets, Markov Models, **System Modeling Language** (**SysML**) based models, and **Unified Modeling Language** (**UML**)-based models were among the most used models in evaluating ACPPS trustworthiness. Some articles even used a combination of different modeling languages for evaluation purposes, moving in the direction of multi-paradigm modeling [21], which includes meta-modeling, abstraction, and multi-formalism.
- Some used domain-specific and graphical modeling formalisms, while others used general-purpose programming languages to construct simulation models for evaluation purposes. These languages were used either for model construction, model-to-model transformations, or in code generation from pre-existing models. A combination of modeling languages was also observed as an approach used to construct models depicting the system's behavior and evaluating it for its trustworthiness. SysML, UML, and Petri Nets were among the most reported modeling languages. All the while, G-code and Python programming languages were mainly reported to have been used for model-based evaluation of ACPPS's trustworthiness.

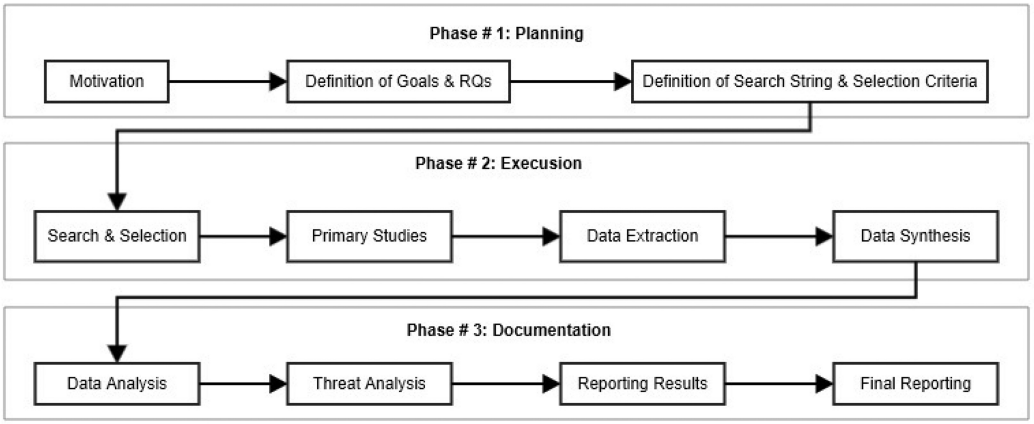

Fig. 2. Phases adopted for review studies.

— The proposed model-based trustworthiness evaluation techniques were tested on various case studies ranging from testbeds, simulations, numerical examples, framework, and data sources to real-world case studies. The case studies included either the entire ACPPS or a component of it to be evaluated.

In what follows is Section 2 describing the research method adopted for this study, Section 3 discussing the threat to validity for this study along with the mitigation strategies adopted, Section 4 presenting the results obtained along with its analysis, Section 5 presenting a discussion of our findings, Section 6 comparing our work with existing related literature, Section 7 concluding this study with final Section 8 presenting potential future works.

2 RESEARCH METHOD

We planned and executed this study using the guidelines by Kitchenham and Brereton, and Garousi et al. on systematic reviews in software engineering [45, 60]. The adopted process consists of three phases, namely: *planning*, *execution*, and *documentation*, as depicted in Figure 2.

Phase # 1: Planning. This phase focused on (i) defining the need to conduct this systematic mapping review on the model-based evaluation of trustworthiness in CPPS, (ii) defining the **research goals** (**RG**) and related **research questions** (**RQs**), and (iii) defining the research protocol to be followed to carry out the study in a systematic manner. The main outcome of this phase is a detailed protocol.

Phase # 2: Execution. This is the phase where all the steps planned during the research protocol definition were performed. These activities were: *search and selection*, *data extraction form definition*, *data extraction*, and *data analysis*.

— Search and selection: We performed an automatic search for peer-reviewed literature on a set of four scientific databases and indexing systems: IEEE Xplore, ACM Digital Libraries, Scopus, and Web of Science. Once the peer-reviewed studies were collected, we filtered them using the selection criteria defined in the planning phase. The filtered studies were then used to perform an exhaustive forward and backward snowballing (using Google Scholar) as suggested by Wohlin et al. [81].

— During the data extraction form definition step, we defined a set of parameters that we used to classify and compare the obtained set of selected primary studies. We did that using the standard key-wording process [70]. The result of this activity was an extraction form.

Table 1. Research Goal

Purpose	Identify, classify, and evaluate
Issue	publication trends, supporting tools, and applications
Object	model-based evaluation of trustworthiness aspects in cyber-physical production systems
Viewpoint	researchers and practitioners.

— During the data extraction step, we analyzed each of the identified primary studies to fill the data extraction forms. We collected and aggregated the filled forms for data analysis and synthesis.
— The goal of the data analysis step was to provide answers to the defined RQs. In the data analysis step, we analyzed the extracted data. By the end of this step, we had performed both quantitative and qualitative analyses of the collected data using vertical and orthogonal analyses.

Documenting phase. The goal of this phase was two folds Analyze and document possible threats to validity that could affect this literature review. Document the procedure of the study along with its findings. To enable the independent replication and verification of this study, we provide a complete and public study replication package containing the data obtained from the search and selection procedure along with the complete list of the primary studies, the data extraction form, and the script used for data analysis and synthesis.[1]

2.1 Research Goal and Questions

The RG of this systematic mapping has been defined as: *identification, classification, and evaluation of publication trends, trustworthiness aspects and attributes being evaluated, modeling formalisms used, and model-based tools used.* This RG has been defined based on Goal-Question-Metric perspectives [13] and is represented in Table 1.

Based on the defined goal, the following are the four RQs contributing to the unique and complementary objectives of this study:

RQ1: What are the main research trends in the model-based evaluation of ACPPS's trustworthiness?
RQ2: Which aspects and attributes of trustworthiness are addressed by current studies?
RQ3: Which are the main modeling languages used for evaluating trustworthiness in ACPPS?
RQ4: What model-based software tools have been used to evaluate trustworthiness in ACPPS?

Answer to *RQ1* provides us with an insight into the kind of research venues that are focusing on our research object *"model-based evaluation of ACPPS's trustworthiness"*. With *RQ2*, we identify which aspects of trustworthiness (such as *trust, dependability, resilience, security, and safety*) have been implemented and evaluated in ACPPS. These aspects of trustworthiness would be the ones reported by the studies' authors. By answering *RQ3*, we draw a picture of the modeling languages used to evaluate ACPPS trustworthiness. Finally, with *RQ4*, we determine the tools used in evaluating ACPPS's trustworthiness via models.

2.2 Search and Selection Strategy

The final set of relevant studies required for this systematic mapping study was collected using the steps depicted in Figure 3. To collect the set of primary studies relevant to this study, we referred

[1]The replication package is available at https://github.com/MaryamZahidMDU/Model-Based_Trustworthiness_Evaluation_ACPPS_replication_package.git

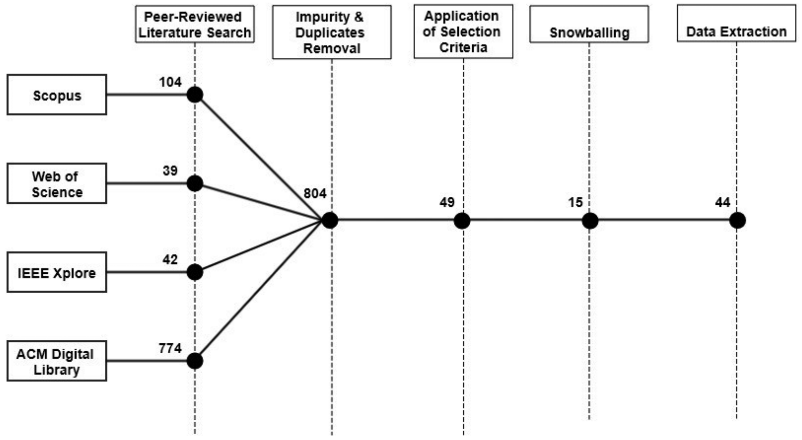

Fig. 3. Overview of the search and selection process.

to four of the largest and most complete scientific databases and indexing systems in the domain of both computer science and software engineering, namely, *SCOPUS*, *Web of Science*, *IEEE Xplore Digital Library*, and *ACM Digital Library* (Table 2). These databases were chosen based on their reputation in terms of being the most effective source of supporting systematic reviews in the domain of both software engineering and computer science, and high accessibility [20, 60].

We started off by creating a search string based on the research goals and questions for exercising the selected databases and indexing systems. Since some of our keywords, such as ACPPS and trustworthiness, are also referred to using other terms, we designed our search string to contain all such possible nuances. Any possible threats to validity associated with the definition of our search string are discussed in Section 3. Furthermore, to avoid coming across a large number of irrelevant articles, we designed the search string to be more focused on our targeted area. The designed search string consisted of five parts summarised in Table 3.

The first part of the search string focused on all possible synonyms of ACPPS, including articles on smart manufacturing processes or smart production systems. The second part of the string focused on aspects of trustworthiness being evaluated in ACPPS, such as trust, dependability, resilience, security, or safety. These attributes were confirmed using the available taxonomy of trustworthiness attributes [40]. Followed by which is the third part of the search string, limiting our search to ACPPS. Here, multiple synonyms were selected to define autonomy in a software system where each synonym covered a particular level of autonomy. The last two components of the search string focus on articles that use model-based techniques for evaluation or analysis purposes. Therefore, the final search string is:

> *("Cyber Physical System" OR CPPS OR CPS) AND ("Production" OR "Manufacturing" OR "Industry 4.0") AND ("Trust" OR "Dependability" OR "Resilience" OR "Security" OR "Safety") AND ("Autonomous" OR "Self-Healing" OR "Intelligent" OR "Self-Sustainable" OR "Self-Repairing" OR "Smart") AND ("Model Based" OR "Model Driven") AND ("Evaluation" OR "Analysis")*

Exercising the defined search string in the selected digital libraries using an automatic search for peer-reviewed literature resulted in a total of *959* potential studies. Removing impurities (articles other than research papers) and duplicates from the obtained initial set of studies reduced the overall size of the dataset to a new total of *804* articles. This reduced dataset still contained articles that lay outside the scope of this systematic mapping study and thus was further scrutinized based

Table 2. Selected Indexing Systems, and Electronic Databases Used

Name	Type	URL
Scopus	Indexing System	http://www.scopus.com
Web of Science	Indexing System	https://www.webofknowledge.com/
IEEE Xplore	Electronic Databases	http://ieeexplore.ieee.org
ACM Digital Libraries	Electronic Databases	http://dl.acm.org

Table 3. Components of Search String

Search String Components	Keywords	Description
ACPPS	("Cyber-Physical System" OR CPPS OR CPS) AND ("Production" OR "Manufacturing" OR "Industry 4.0")	This component focuses on the keywords related to Cyber-Physical Production Systems
Trustworthiness	("Trust" OR "Dependability" OR "Resilience" OR "Security" OR "Safety")	This part of the search string focuses on the aspects of trustworthiness being evaluated in ACPPS
Autonomous	("Autonomous" OR "Self-Healing" OR "Intelligent" OR "Self-Sustainable" OR "Self-Repairing" OR "Smart")	These were the acronyms used to focus on the AI aspect of the ACPPS
Model	("Model-Based" OR "Model Driven")	These keywords were used to extract articles focusing on model-based techniques used for evaluation
Evaluation	("Evaluation" OR "Analysis")	These were the possible keywords against "evaluation" we used to construct our final search string

on the selection process proposed by Ali and Petersen [6]. The inclusion and exclusion criteria used to ensure an objective selection of primary studies are as below:

— Inclusion criteria:

(1) Contents of the article focus on the model-based evaluation of trustworthiness in ACPPS.
(2) Articles written in English.
(3) Articles that are peer reviewed [82] and not published as technical reports, white articles, and editorial articles.
(4) Type of article is either a journal article, a book chapter, a conference paper, or a workshop article.
(5) Articles that do not only present a vision or a viewpoint, or report keynotes, discussions, opinions, editorials comments, prefaces, tutorials and anecdote articles, and presentations in the form of slides without any associated research articles.

— Exclusion criteria:

(1) Articles that discuss the model-based evaluation of autonomous CPS's trustworthiness as background only or as a side topic.
(2) Articles that focus on CPS other than the cyber-physical production or manufacturing systems (i.e., smart-factory applications within Industry 4.0).
(3) Articles focusing on the model-based evaluation of autonomous CPS properties unrelated to trustworthiness.

Table 4. Data Extraction Form

Facets	Cluster	Category	Description	Value
RQ1	Publication Details	Publication Title	Identifies the title of the primary studies	String
		Authors	Highlights the names of the author	String
		Venue	Identifies the title of the venues	String
		Venue Type	Highlights the venue type	String
		Year	Highlights the publication year	Numeric
RQ2	Aspects of Trustworthiness	–	Identifies the aspect of trustworthiness implemented and evaluated according to the taxonomy in [65]	String
RQ3	Modeling Languages	–	Identifies the modeling languages used for evaluating ACPPS's trustworthiness [12]	String
RQ4	Tools	–	Identifies the developed model-based evaluation tools	String

(4) Articles that do focus on evaluating autonomous CPS's trustworthiness but not using model-based evaluation techniques.
(5) Articles that present secondary or tertiary studies.

To undergo the next step of the execution phase, a particular study must meet all the inclusion criteria and none of the defined exclusion criteria. This criteria-based selection process led to a new set of selected studies consisting of a total of *49* articles that can potentially be declared as primary studies. Later, during the execution phase, a closed recursive backward and forward snowballing activity was performed [81], thus minimizing potential bias concerning construct validity [49], all the while resulting in *15* additional peer-reviewed articles leading to a final set of *44* primary studies.

2.3 Data Extraction

A data extraction form was created to further proceed with the extraction and data collection process from the final set of primary studies. Table 4 depicts the extraction form that is composed of four facets, each targeting an individual RQ. The first facet focuses on the standard information regarding all the peer-reviewed literature, such as title, authors, and other publication details. For the rest of the targeted RQs RQ2–RQ4, we adopted a *keyword based* systematic process to develop the extraction form responsible for considering the characteristics of the selected primary set of studies [70]. We started off by reading the full text of the primary studies. This resulted in a collection of keywords and concepts obtained from the studies. These keywords and concepts were then clustered and categorized based on a process similar to the one defined as a sorting mechanism of the grounded theory methodology [25]. In case of reaching a point where the collected information, although relevant, was not captured by the current extraction form, the information was reviewed. If the inclusion of such information was deemed necessary, the form was refined to accommodate that piece of information but only after re-analysis based on the modifications made to the data extraction form. During this phase, 18 more peer-reviewed articles were excluded, leading to a new total of *44* articles in the final set of primary studies. These articles were excluded from the final set only after a full-text analysis of the studies. Table 5 represents the final set of primary studies collected after analysis and extraction.

2.4 Data Analysis and Synthesis

With the help of guidelines provided by Cruzes et al. [31], the extracted data from the final dataset obtained (Table 5) was collected, analyzed, and synthesized to understand and classify the current

Table 5. Primary Studies

ID	Title	Authors	Year	Ref.
P1	Lightweight Mutual Authentication and Privacy-Preservation Scheme for Intelligent Wearable Devices in Industrial-CPS	Jan et al.	2022	[51]
P2	On the impact of empirical attack models targeting marine transportation	Bou-Harb et al.	2020	[18]
P3	A New Threat Intelligence Scheme for Safeguarding Industry 4.0 Systems	Moustafa et al.	2019	[67]
P4	Trustworthiness Modeling and Analysis of Cyber-physical Manufacturing Systems	Yu et al.	2016	[86]
P5	Dynamic Probabilistic Model Checking for Sensor Validation in Industry 4.0 Applications	Xin et al.	2020	[84]
P6	Orthogonal Uncertainty Modeling in the Engineering of Cyber-Physical Systems	Bandyszak et al.	2020	[10]
P7	A model-based approach to qualified process automation for anomaly detection and treatment	Chen et al.	2016	[26]
P8	Fostering concurrent engineering of cyber-physical systems a proposal for an ontological context framework	Duan et al.	2016	[32]
P9	Integrated tool chain for model-based design of Cyber-Physical Systems: The INTO-CPS project	Larsen et al.	2016	[62]
P10	Optimal replacement model for the physical component of safety critical smart-world CPSs	Alemayehu et al.	2021	[5]
P11	Towards Model-Based Performability Evaluation of Production Systems	Bucaioni et al.	2020	[21]
P12	Use case-based consideration of safety and security in cyber physical production systems applied to a collaborative robot system	Lichte et al.	2018	[63]
P13	An Emerging Industrial Business Model considering Sustainability Evaluation and using Cyber Physical System Technology and Modeling Techniques	Watanabe et al.	2016	[79]
P14	A simulation-based platform for assessing the impact of cyber-threats on smart manufacturing systems	Bracho et al.	2018	[19]
P15	Self-adaptive traffic control model with Behavior Trees and Reinforcement Learning for AGV in Industry 4.0	Hao et al.	2021	[50]
P16	A Smart Maintenance tool for a safe Electric Arc Furnace	Fumagalli et al.	2016	[43]
P17	SAVE: Security & safety by model-based systems engineering on the example of automotive industry	Jasp et al.	2021	[52]
P18	Simulation-As-A-Service: A simulation platform for cyber-physical systems	Treichel et al.	2021	[78]
P19	An Architecture-based Modeling Approach Using Data Flows for Zone Concepts in Industry 4.0	Kern et al.	2020	[57]
P20	Resilient fault diagnosis under imperfect observations-A need for Industry 4.0 era	White et al.	2020	[80]
P21	GAN-Sec: Generative Adversarial Network Modeling for the Security Analysis of Cyber-Physical Production Systems	Chhetri et al.	2019	[28]
P22	5-dimensional definition for a manufacturing digital twin	Bazaz et al.	2019	[14]
P23	Finding dependencies between cyber-physical domains for security testing of industrial control systems	Castellanos et al.	2018	[23]
P24	Robust Digital Twin Compositions for Industry 4.0 Smart Manufacturing Systems	Preiveneers et al.	2018	[71]

(Continued)

Table 5. Continued from previous page

ID	Title	Authors	Year	Ref.
P25	Enhancing cyber-physical security in manufacturing through game-theoretic analysis	Desmit et al.	2018	[35]
P26	Security viewpoint in a reference architecture model for cyber-physical production systems	Ma et al.	2017	[9]
P27	Multi-Scale Software Network Model for Software Safety of the Intended Functionality	Zhitao et al.	2021	[83]
P28	Buoy Sensor Cyberattack Detection in Offshore Petroleum Cyber-Physical Systems	Lin et al.	2020	[68]
P29	Enhancing IIoT networks protection: A robust security model for attack detection in Internet Industrial Control Systems	Khan et al.	2022	[59]
P30	Novel model for boosting security strength and energy efficiency in internet-of-things using multi-staged game.	Ambore et al.	2019	[7]
P31	Cyber-Physical Security Evaluation in Manufacturing Systems with a Bayesian Game Model	AliReza et al.	2020	[87]
P32	Model-based documentation of context uncertainty for cyber-physical systems	Torsten et al.	2018	[11]
P33	Enabling model testing of cyber-physical systems	González et al.	2018	[48]
P34	Cybersecurity analysis of smart manufacturing system using game theory approach and quantal response equilibrium	AliReza et al.	2018	[88]
P35	Attack path analysis for cyber physical systems	Georgios et al.	2020	[56]
P36	Cyber-physical vulnerability assessment in manufacturing systems	DeSmit et al.	2016	[33]
P37	An S4PR class Petrí net supervisor for manufacturing system	MH et al.	2016	[2]
P38	A game-theoretic approach to model and quantify the security of cyber-physical systems	Orojloo H et al.	2017	[69]
P39	An approach to cyber-physical vulnerability assessment for intelligent manufacturing systems	DeSmit et al.	2017	[34]
P40	Modeling availability risks of IT threats in smart factory networks–a modular Petri net approach	Berger et al.	2019	[16]
P41	A Semantic Framework with Humans in the Loop for Vulnerability-Assessment in Cyber-Physical Production Systems	Jiang et al.	2019	[54]
P42	IT Availability Risks in Smart Factory Networks - Analyzing the Effects of IT Threats on Production Processes Using Petri Nets	Berger et al.	2022	[17]
P43	Model-based Stochastic Error Propagation Analysis for Cyber-Physical Systems	Fabarisov et al.	2020	[37]
P44	Threat modeling for industrial cyber physical systems in the era of smart manufacturing	Jbair et al.	2022	[53]

state-of-the-art in the domain of *model-based trustworthiness evaluating of ACPPS* [61]. This study presents a quantitative analysis of the state-of-the-art discussed in the primary studies using the vertical analysis technique [41]. This type of analysis helped find and collect information on each of the defined facets of the data extraction form using the line of argument synthesis process [82]. During this process, each primary study was individually analyzed to have its features classified based on the parameters defined in the data extraction form.

3 THREATS TO VALIDITY

We conducted this study using a set of well-established guidelines proposed for systematic studies in software engineering. However, there still exists the possibility of certain validity threats

affecting the results of our research. Considering this, we describe the main validity threats and their associated mitigation strategies below.

3.1 Threats to External Validity

One possible threat to external validity could be the existence of other terms used in place of CPPS, automation, trustworthiness, and evaluation, which might limit our search coverage. To mitigate this threat, we expanded our search string to contain alternative terms related to the ones mentioned above. When selecting the studies, we considered only those articles published in English. Excluding articles published in another language may affect the validity of this study; we believe the impact of such a threat to be minimal as English is the *de-facto* standard language for almost all scientific documents, especially in the domain of computer science and software engineering.

3.2 Threats to Construct Validity

To mitigate the threats to the construct validity of having a single source, we exercised four different scientific databases and indexing systems among the most complete and reputable databases in the field of software engineering. To mitigate these threats further, we performed a closed recursive backward and forward snowballing. The study selection was performed using a set of selection criteria.

3.3 Threats to Internal Validity

To mitigate threats to internal validity, we followed a set of well-established guidelines proposed for systematic studies in software engineering. Furthermore, we conducted a cross-analysis between the different categories of extraction form and performed sanity checks on the extracted data to maintain consistency of the extracted data.

3.4 Threats to Conclusion Validity

The well-defined processes of this study were constantly and systematically applied and documented. A complete and public replication package allowing the reproduction and verification of our study has been made available. All authors participated in the planning, execution, and documentation phase of the study. Moreover, an extraction form was built using well-established taxonomies and collecting values emerging from the set of primary studies.

4 RESULTS

This section presents our findings. Each study in the final set of selected studies was analyzed individually using the designed data extraction form to have its main features classified. The whole set of primary studies was further analyzed as a whole to identify any other potential patterns, trends, and research gaps. While analyzing primary studies against each category of the data extraction form, multiple values were extracted if available. All the while, no values were extracted for the studies that did not provide any values. Thus, the total occurrences reported in the plotted graphs may vary from the total number of primary studies.

4.1 Publication Trends (RQ1)

This RQ was designed to identify the publication trends in the domain of model-based evaluation of ACPPS's trustworthiness. The answer to this question was collected regarding publication year, venues, sources of publications, and type of article published. It is observed from the collected data that the domain is relatively new. Although many studies have been conducted on the evaluation of trustworthiness, they are either focused on generic IT systems or CPS, or the evaluation process

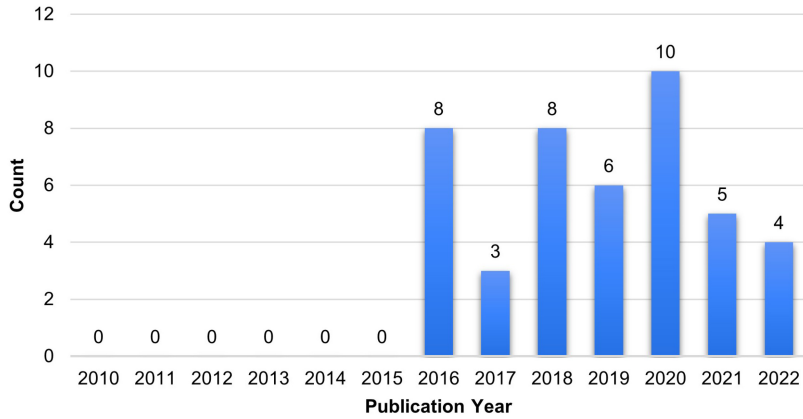

Fig. 4. Number of publications over time (year 2022 being incomplete).

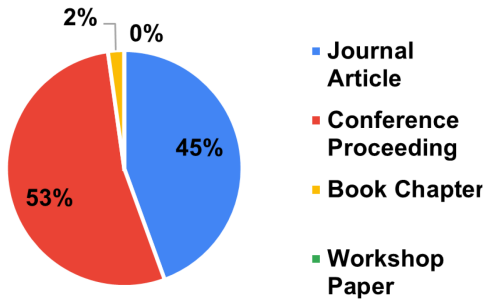

Fig. 5. Type of published articles.

itself is not model-based in nature. Figure 4 shows a gradual rise in the number of publications in the targeted domain, starting from the year 2016 and reaching a peak in the year 2020 with a percentage of almost *22.2%* of the relevant articles published in the year. This is also where we begin to see a gradual decline in the number of publications in the domain, reaching a minimum of *8.9%* of the total number of articles in the year 2022. Observing the publication trend, we keep seeing a decline in the number of publications in the years 2017, 2019,2021, and 2022, with 2017 being the year with the least number of publications. All the while, a minor increase in the number of publications can be found in the year 2018, having a total of *8* articles published in the year before reaching a peak in the year 2020 with a total of **10** articles being published. As for the decline in the number of publications in the years 2021-present, a reason can be the fact that certain articles are not being included as they have not been published yet and thus made available in the targeted digital libraries or indexing systems.

Among the various sources of these publications, most of the articles were available in Scopus *(15)*, followed by IEEE Xplore *(14)* and Science Direct *(12)*. On the other hand, AIS library and ACM digital libraries were among the sources with the least number of publications in the domain *(1 and 0 articles, respectively)*.

The most frequent type of articles in the domain we came across were conference papers *(53%)* followed by journal articles *44%* (Figure 5). Although the automated search on digital libraries did yield several workshop papers, none passed our selection criteria, thus having the percentage of *0%* in our final set of primary studies. As for the book chapters, only one was selected as a primary

study based on the inclusion and exclusion criteria designed for this systematic mapping study. During the search and selection process, no particular conference venue stood out as the main venue of the selected set of primary studies. The research community did not prefer a single venue to have their studies on model-based evaluation of ACPPS's trustworthiness published. However, conferences on AI and model-driven engineering seemed to have been the venue preferred for most of these articles. All the while, journals such as *"Computer in Industry"* and *"IEEE Transactions on Industrial Informatics"* were the most targeted journals to have articles published that are of relevance to our studies.

Highlights – RQ1 Recent Trends in Publications

- The earliest published work on the model-based evaluation of ACPPS's trustworthiness started appearing in 2016.
- The popularity of the domain reached its peak in the year 2020, followed by a gradual decline.
- Conferences on AI and model-driven engineering, *"Computer in Industry"* journal and *"IEEE Transactions on Industrial Informatics"* journal were among the most targeted venue.
- Scopus, IEEE Xplore, and ACM Digital Library were the main sources of articles in the targeted domain.

4.2 Aspects and Attributes of Trustworthiness (RQ2)

Trustworthiness in a system is a combination of the following attributes that collectively help ensure user's and stakeholder's trust in the system [1, 27, 42, 64]:

— *Performability:* a measure to evaluate the system's performance regarding how well it executes its functionalities under normal and abnormal conditions.
— *Integrity:* a property of the data or the information being manipulated by the system, such that it is accurate, consistent, complete, and trustworthy.
— *Confidentiality:* an attribute associated with the system's security and privacy, such that the information being manipulated is accessible and usable only by the authorized users or processes of the system. It also implies that the data or the information being used is protected from unauthorized disclosure or interception.
— *Availability:* a characteristic of data or the information being used by the system such that it is accessible and usable only by the authorized users or processes of the system as per need, all the while being protected from unauthorized denial of service or disruption.
— *Reliability:* can be defined as the probability of the system executing a set of specified functionalities without failure under certain conditions and time periods.
— *Maintainability:* a property associated with the software system's design, complexity, and documentation in terms of ease in modifying, repairing, or enhancing the system's functionalities to meet the changing requirements or correct the identified defects.
— *Scalability:* an attribute related to the resources, architecture, and configuration of the system that can be defined as the capability of the system to handle the varying workloads or demands without compromising its overall performance, availability, reliability, and the quality of service.
— *Robustness:* is a property of the system related to its resilience and fault-tolerance, concerned with its ability to deal with occurring errors, faults, unexpected inputs, or conditions without failing operations or producing incorrect outputs.

— *Safety:* is a quality of the system that does not threaten its users, property, or environment under both normal and abnormal conditions.
— *Holistic Security Metrics:* an approach considering all aspects of security in a complete and coherent manner aiming to accomplish the system's security objectives while balancing its other qualities and the needs of its stakeholders.
— *Holistic Resilience Metrics:* an approach considering all aspects of resilience completely and coherently aiming to optimize the system's ability to withstand and recover from adverse conditions or events while maintaining its basic functionalities and other performance factors.
— *Holistic Dependability Metrics:* an approach considering all dependability aspects completely and coherently. The aim is to ensure that the system operates and provides its services as per the users' and stakeholders' needs under various conditions.
— *Holistic Trust Metrics:* an approach considering all the aspects of trust completely and coherently aiming to ensure trust relationships among its different entities while being in a complex and dynamic environment.

This RQ aimed at identifying the reported aspects of trustworthiness evaluated in ACPPS using model-based evaluation techniques. According to the data collected, safety and security as a whole have been the most reported aspects of trustworthiness being evaluated in ACPPS over the last decade, with a percentage of *19%* and *18%*, respectively. Authors of such articles seem to prioritize safety and security above other attributes [65] due to their criticality in industrial settings. According to the author of reference, [65], safety and security are interrelated such that the system's compromised security may lead to an increase in safety-related risks in ACPPS [52]. Looking at these two reported attributes individually; the safety aspect marginally supersedes security as a whole. But when considering the individual aspects within holistic security such as **confidentiality, integrity, and availability** (**CIA**) together (*6%, 7%, and 11%*, respectively) with security as a whole, safety attribute comes second to it in terms of the reported attributes evaluated in ACPPS. Among other aspects of trustworthiness, reliability has also been of keen interest to the researchers, with around *10.5%* of the studies reporting it being evaluated, while scalability, performability, and maintainability were among the least reported attributes evaluated. Furthermore, the results revealed around *8.5%* of the studies focusing on a holistic integrated and holistic evaluation of trustworthiness in ACPPS. Whereas Figure 6 suggests that among the aspects observed to have been reported for evaluation using model-based techniques in the final set of primary studies, dependability and resilience as integrated concepts based on holistic metrics were the least focused on (only *9%* and *6%* of the final set of studies, respectively).

Highlights – RQ2 Focused Aspects of Trustworthiness

- Security as a holistic metric and as its individual indicators, along with safety-related attributes, were among the most evaluated aspects within ACPPS using model-based techniques.
- The results obtained show an opportunity to explore more holistic ACPPS dependability metrics and resilience indicators, such as performability-related ones, as they are, at the moment, the least focused aspects being evaluated using model-based techniques, while the integration of possibly competing trustworthiness attributes is the key to achieve the right implementation tradeoffs (e.g., performance vs. dependability).

4.3 Modeling Languages (RQ3)

The answer to this RQ provides us with an overview of the modeling languages used for evaluating trustworthiness in ACPPS. According to our investigation, a total of 10 different types of modeling languages were identified to have been used to represent the structural and behavioral aspects of

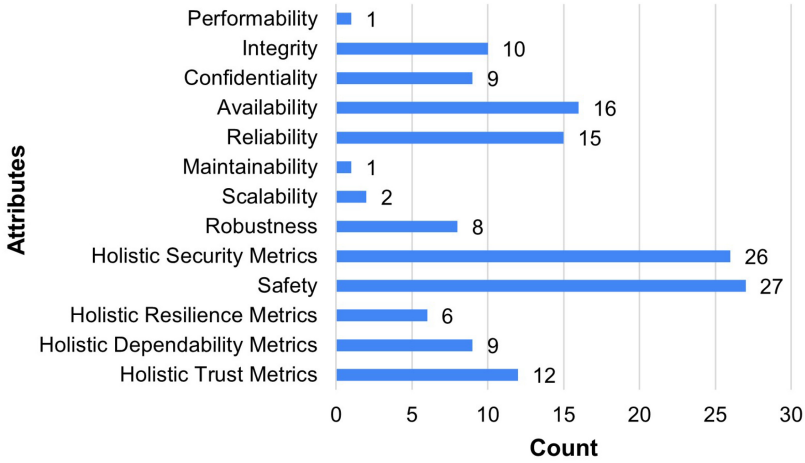

Fig. 6. Attributes of ACPPS trustworthiness evaluated through models in the reviewed articles according to authors' reported focus, taxonomy, and keywords.

Table 6. Modeling Languages Used

modeling Languages	**IDs**
Petri Nets	P4, P11, P13, P37, P40, P42
SySML	P2, P9, P12, P17, P20, P33, P43
UML	P8, P18, P26, P41, P43
OUM	P6, P32
RAAML	P16
EAST-ADL	P7
AADL	P43
Other GML	P14, P15, P27, P35
Other BPMN	P24
Other DSL	P11, P19

ACPPS to be evaluated for its trustworthiness. The results presented in Table 6 show the use of **System modeling Language** (**SysML**) the most for constructing models of the underlined embedded systems, and Petri Nets capturing the behavior of the system to be evaluated with a frequency of *23%* each. Another most commonly used modeling language identified is the **Unified Modeling Language** (**UML**) with a frequency of *16.1%*, followed by **Graph modeling Language** (**GML**) for representing interactions between various components of the system using graphs and trees as an instance (12.9%). Some showed interest in the use of **Orthogonal Uncertainty Modeling** (**OUM**) language (*6.7%*) to represent the identified uncertainties in a system to be evaluated [10, 11], and **Risk Analysis and Assessment Modeling Language** (**RAAML**) (*6.7%*) representing fault trees constructed for evaluation of trustworthiness related aspects in the system. At the same time, some preferred developing their own Domain-specific modeling language to construct a Logical and Functional Layered Architecture of the system [57].

Some articles presented the use of a combination of ontologies along with UML (*6.7%*) to allow context modeling of the system via component models to have the system evaluated using model-checking tools. Since Markov Models, especially **Hidden Markov Models** (**HMM**), are

language-independent, no specific modeling or programming language has been stated to have been used for evaluating the trustworthiness of ACPPS.

Highlights – RQ3 Model Language

- A total of 10 different modeling languages were reported to have been used for evaluating the trustworthiness of ACPPS.
- Petri Nets, SysML, and UML were among the most reported modeling languages used.
- Articles using HMM did not specify any language used.
- Some used a combination of modeling languages to represent both the structural and behavioral aspects of ACPPS to be evaluated.
- Some used instances of models specified by the safety analysis standards specifically designed for that domain.

4.4 Tools (RQ4)

By answering this RQ, we provide a snapshot of the model-based tools used for evaluating the trustworthiness of ACPPS using models. The results obtained show a total of *42* different tools used for evaluation purposes consisting of either preexisting tools used for model constructions or model-based evaluation itself or the tools developed explicitly for evaluating a particular ACPPS system itself (*7%*), a simulation, or a specific component. Among these *32*, different model-based tools were reported to have been used for evaluation purposes. These tools allowed evaluation of trustworthiness aspects mainly at the design level, focusing on the system as a whole using system models, a component using component models, or other models focusing on a particular layer of the ACPPS's software and hardware architecture. In some cases, no particular tool was reported to have been used for model-based evaluation of ACPPS's trustworthiness aspects. Some studies also suggested using a combination of model-based tools to generate models to be evaluated for the system's trustworthiness. Going into the depth of the results obtained, Petri Net tools (*21.8%*) such as IOPT-Tool [79], and **Colored Petri Nets** (**CPN**) Tool [16] were the most reported model-based tools used for evaluating the trustworthiness of ACPPS. Followed by this came MATLAB [69] (*12.5%*) used specifically for model-based evaluation of the trustworthiness of ACPPS using game theoretic models. Some have also opted for developing metamodels using modeling tools such as CAD *12.5%* [34] to help develop models depicting the system's behavior to be evaluated for its trustworthiness. Figure 7 represents a statistical analysis of system evaluation tools based on modeling concepts such as Markov Model Chains and other existing simulations. For example, a study suggested using a FERAL Simulator to model the system and its interactions with other subsystems before evaluating it [78]. Simultaneously, a keen interest of the research community and the practitioners can also be seen in the use of self-developed tools evaluating such autonomous systems (*9.3%* each).

Highlights – RQ4 Model-based Tools

- The model-based tools reported can be categorized as either those used for modeling the component of ACPPS to be evaluated or those that were purely used for evaluating the system.
- Petri Net tools were among the most reported tools used for evaluation purposes, followed by MATLAB and CAD used to implement game theoretic models.
- Authors of some works developed domain-specific tools to evaluate the system's trustworthiness.

5 DISCUSSION

Evaluation of ACPPS trustworthiness using model-based techniques started to gain popularity after 2020, with a total of *19* primary studies published in the duration. Due to the novelty of the

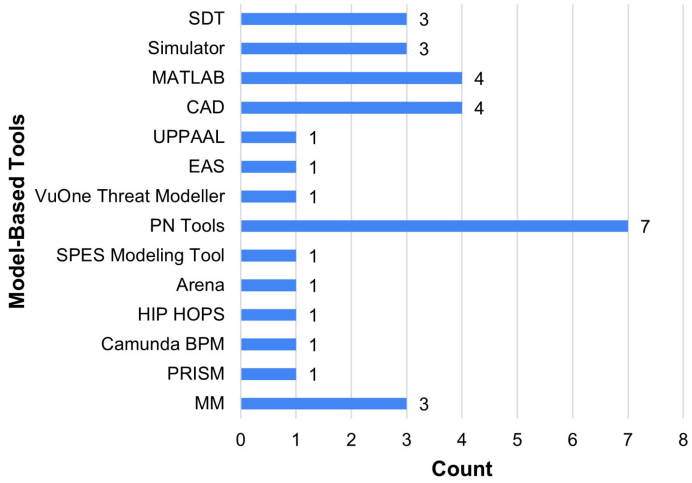

Fig. 7. Model-based tools used for evaluation of ACPPS's trustworthiness.

domain, not all aspects of ACPPS trustworthiness were covered to the full extent in the works surveyed in this study. In general, most of the proposed or discussed model-based trustworthiness evaluation techniques were focused on the evaluation of components of a system being designed rather than the entirety of the system.

The trustworthiness attributes evaluated consisted of the holistic metrics such as holistic security, dependability, and resilience metrics and indicators of these holistic metrics such as CIA (indicators of holistic security attribute). The studies evaluating holistic metrics focused on generic risks associated with those particular metrics.

For evaluation purposes, where the majority of the primary studies presented the use of model-based evaluation techniques, some opted for using non model-based techniques combined with models to evaluate the system's aspects of trustworthiness. One such example was the usage of ontologies with OUM [10, 11]. Furthermore, not all languages used for model construction were modeling languages; rather, some used programming languages to construct models (such as the HMM model and the Attack Trees), meta-models (as Seen in Figure 8), or for storing the constructed model of the system for model-to-model transformations [36], evaluation purposes, or code generation. Articles using G-code mainly used graph-based techniques for trustworthiness evaluation. One example is the use of Intersection mapping for vulnerability detection and decision trees for assessing the identified vulnerabilities [33]. Aspects analyzed using G-code included Loss of Information, Data Inconsistencies, Time Until Detection, Relative Frequency, and Lack of Maturity [34]. Studies based on Game Theoretic Models used both Bayes–Nash Equilibrium and Quantal Response Equilibrium to capture the interactions taking place between the attacker and the defender in a given smart manufacturing system [87].

Statistically analyzing all the reported languages used for evaluating trustworthiness-related attributes, XML, G-Code, Python, Fortran, and Ontologies emerged with a frequency of *4.5%*, *9.1%*, *4.5%*, *2.3%*, and *4.5%*, respectively. Whereas when limiting the statistical analysis of reported languages used for evaluation purposes to only non-model-based languages, these frequencies changed to *18%*, *37%*, *18%*, *9%*, and *18%*, respectively (as seen in Figure 8). The articles reporting the use of such languages can be seen in Table 7.

Authors reporting the use of non-model-based languages in most cases evaluated trustworthiness-related attributes such as reliability, availability, holistic security, resilience,

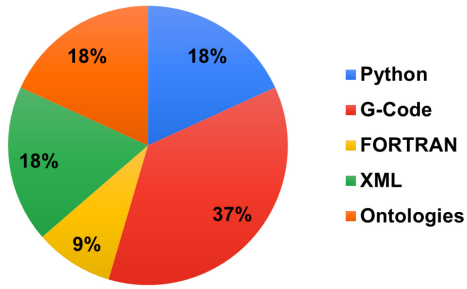

Fig. 8. Non-model-based languages used for evaluating trustworthiness in ACPPS.

Table 7. Non-model-based Languages Used

Non-Model Based Languages	IDs
Python	P28, P29
G-Code	P39, P36, P21, P25
Fortran	P28
XML	P22, P23
Ontologies	P41, P8

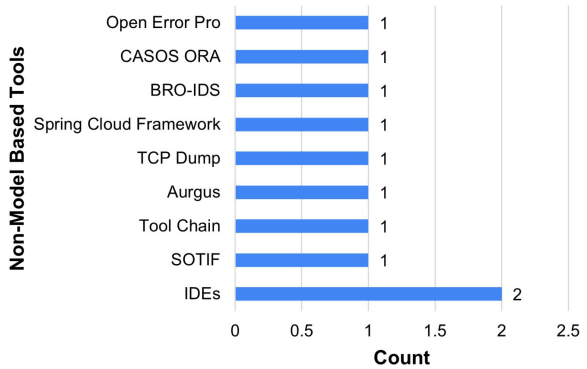

Fig. 9. Non-model-based tools used for evaluation of ACPPS's trustworthiness.

and trust via G-Code implementation on CAD/CAM tools, while others used XML, Python or Fortran code implementations on IDEs mainly to evaluate safety and reliability.

Looking into the tools used for evaluation purposes, not all tools were purely used for evaluation purposes; rather, some were used for model construction or data collection required to assist in the model-based evaluation of trustworthiness aspects. Some of the reported tools used were not even model-based in nature, e.g., **Transmission Control Protocol** (**TCP**) Dump, a tool to capture data packets within a network was used to identify the flow of data between the various components of the system, or the IDEs used to construct meta-models and the models originally required for evaluation purposes. Figure 9 presents all the non-model-based tools used to assist in generating models required for evaluation purposes.

According to Figure 9, IDEs (*20%*) were the most reported tool used to implement the code written in non-model-based languages. The rest of the tools reported in the selected set of primary studies appeared with the frequency of *10%*. The reason for such an anomaly is that either

these tools were mainly used in conjunction with other model-based tools, e.g., the use of CAD tool with CAM tool [33], or as discussed earlier was used to extract the required data for analysis or for implementing a meta-model only before constructing the actual model evaluating the trustworthiness-related attribute.

An integrated analysis, as seen in Table 8, presents a list of all the tools used for model construction and thus evaluation the respective trustworthiness-related attribute. Looking at only the attributes being evaluated, most reported were either safety, availability, reliability, or holistic security metrics, for which SysML was the most used modeling language implemented using tools such as MATLAB or, in other cases, a self-developed tool based on the stakeholder's requirements. On the other hand, considering the modeling language alone, the most frequently reported modeling language used for evaluating the attributes was Petri Nets [22]; for which most of the authors opted for Petri Net Toolbox or MATLAB, evaluating almost all of the attributes except for maintainability, and holistic resilience metrics. Although only four articles reported the use of GML, the integrated analysis reveals a significant use of GML for evaluating almost all the trustworthiness-related attributes except for maintainability, scalability, and performability. In this case, the authors either preferred the use of MATLAB or opted for a self-developed tool.

Further analyzing the obtained results for reported modeling language used for the attribute evaluated, RAAML was the least reported language being implemented using a simulation tool (Smart Water Monitoring (Web-based Tool)) responsible for assessing the safety and reliability of the system in question. Authors reporting the use of DSL implemented them using their own developed tools and, therefore, were among those reported the least number of times for evaluating the attributes. Overall, the use of both qualitative and quantitative assessment tools has been made to evaluate the attributes. One observation made from this study is the lack of focus on evaluating certain attributes such as maintainability, scalability, performability, and holistic resilience metrics, and therefore have the least associated modeling languages reported.

As for the case studies adopted for demonstrating the working of the model-based evaluation technique, a total of *11* different case studies were reported to have been used. Most of which were either generic case studies or were related to smart grids. Most of the studies used either a simulation of a smart **Industrial Internet of Things** (**IIoT**) involving the entire manufacturing process or case studies on the transportation of goods or materials required for the production process, e.g., platooning application or a conveyor belt, with a percentage of *24.4%* each. According to the statistics obtained, the second most used type of case studies (*22%*) were the real-world case studies on ICS to implement the proposed model-based trustworthiness evaluation technique. Since some of the articles focused on evaluating the trustworthiness of a particular component of ACPPS rather than the entire system, we can see from Figure 10 the use of case studies related to the networking layer of ACPPS, the resource allocation aspect of ACPPS, or the Buoy sensors (*2.4%* each). In some cases, numerical examples, test beds of a particular ICS, and pre-existing globally available data sources (*5%* each) or frameworks such as the Design Science Research Evaluation Framework were also used as a case study (*2.4%*).

Overall, based on our investigation, not all attributes of trustworthiness have yet been evaluated for ACPPS, especially using model-based techniques. Although the studies do present the evaluation conducted, they do not specify the limitations of their evaluation techniques, such as the complexity of their technique, scalability, or generalizability. The proposed techniques seem to focus mostly on one particular layer of the ACPPS architecture, thus ignoring the risks associated with the rest of the architecture, resulting in an incomplete system evaluation. The majority of the studies focused more on risk identification rather than analysis. Also, the techniques proposed are not continuous or automated. Moreover, no method for traceability has been discussed in the articles to help monitor and manage the risks identified during the evaluation process. Therefore,

Table 8. Integrated Analysis

Trustworthiness Attribute	Modeling Language Used	Tool Used
Integrity	Petri Nets, SysML, UML, GML	MATLAB, Petri Nets Toolbox, IOPT, SecuriCAD, FERAL simulator, UPPAAL
Confidentiality	Petri Nets, SysML, GML	MATLAB, Petri Nets Toolbox, SecuriCAD, FERAL simulator, UPPAAL
Availability	Petri Nets, DSL, UML, GML, EAST ADL	MATLAB, Petri Nets Toolbox, IOPT, SecuriCAD, FERAL simulator, HiP-HOPS (FTA), MetaEdit, Camunda BPM, MM Tools
Reliability	Petri Nets, SysML, UML, GML, BPMN, RAAML	MATLAB, Petri Nets Toolbox, IOPT, PRISM, EAS, Self Made Tool, Smart Water Monitoring (Web-based Tool), Arena® from Rockwell Automation, Discrete-Time Markov Chains (DTMCs), probabilistic Model Checker, MM Tools
Maintainability	Self Developed DSL	Self Made Tool
Scalability	UML, Petri Nets	FERAL simulator, Petri Net Toolbox
Robustness	Petri Nets, OUM, SysML, BPMN, GML, EAST ADL	Petri Nets Toolbox, Self Developed Model, Self-Made Tool "Diagnoser", HiP-HOPS (FTA), MetaEdit, Camunda BPM, MM Toolbox
Safety	Petri Nets, EAST ADL, OUM, SysML, UML, BPMN, DSL, GML, RAAML	SPES Modeling Tool, HRD, Arena® from Rockwell Automation, MATLAB, IOPT, Self-Made Tool "Diagnoser", HiP-HOPS (FTA), MetaEdit, Camunda BPM, Self-Made DSL Based Tool, PRISM, Siemens Tecnomatix plant simulation, Smart Water Monitoring (Web-based Tool), VueOne Threat Modeller
Performability	Petri Nets	Petri Net Toolbox
Holistic Security Metrics	Petri Nets, OUM, SysML, UML, GML, DSL	MATLAB, Petri Nets Toolbox, Arena® from Rockwell Automation, EAS, MATLAB, IOPT, Self-Made Tool "Diagnoser", Self-Developed Model, Self-Made DSL Based Tool, PRISM, Siemens Tecnomatix plant simulation, MM Toolbox, VueOne Threat Modeller
Holistic Resilience Metrics	GML, SysML	MATLAB, Self-Made Tool "Diagnoser"
Holistic Dependability Metrics	Petri Nets, GML, SysML, EAST ADL	MATLAB, Petri Nets Toolbox, Tool Chain, HiP-HOPS (FTA), MetaEdit, Camunda BPM, MM Toolbox, PRISM, VueOne Threat Modeller
Holistic Trust Metrics	Petri Nets, GML, SysML, UML	VueOne Threat Modeller, MM Toolbox, FERAL Simulator, Self-Developed Model, Petri Net Toolbox, DTMC, Probabilistic Model Checker, PRISM, MATLAB, EAS

this leads us to factors that need further investigation to help better understand and ensure trustworthiness in ACPPS.

6 RELATED WORK

A CPS integrates embedded computers and communication technologies to various physical domains, including the production domain, thus resulting in a CPPS [47]. Much research has been done on the evaluation of CPPS's performance and reliability over the last decade. However, since

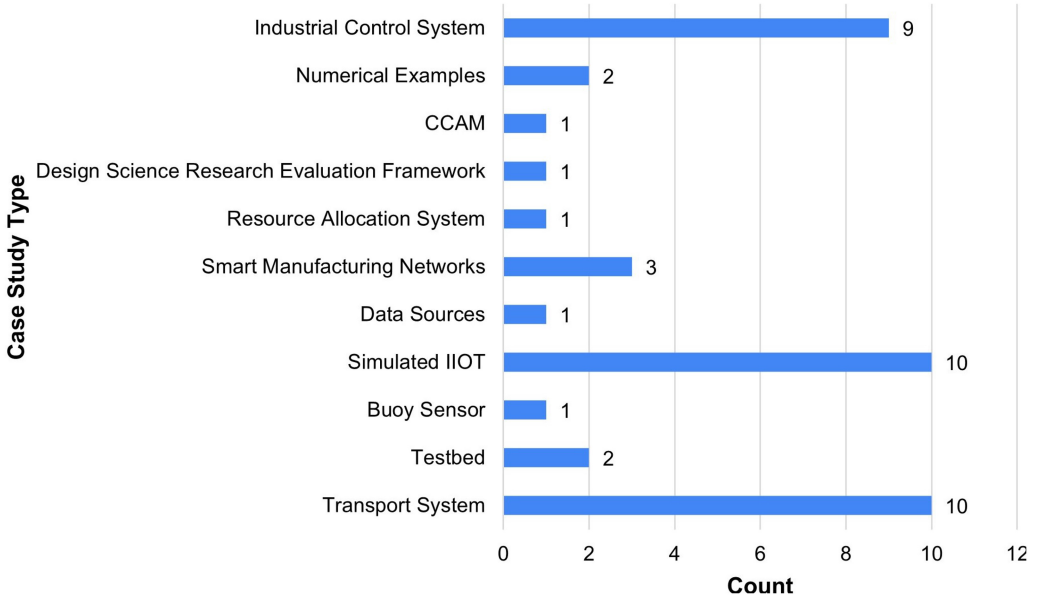

Fig. 10. Case studies used for demonstration.

ACPPS is a relatively new domain with artificial intelligence integrated into traditional CPPS [4], the domain is still being explored. To the best of our knowledge, the work presented in this article is the first systematic mapping study presenting an analysis of the model-based evaluation of trustworthiness implemented in ACPPS, highlighting the aspects of trustworthiness evaluated, modeling formalism used, tools used, and the case studies used to present the evaluation process. Nonetheless, several studies have been conducted on evaluating trustworthiness in CPPS.

Jairo et al. presented a survey of surveys on the security and privacy defenses (prevention, detection, and responses) in CPS such as manufacturing systems and ICS, and so on [47]. The article discusses the use of detection algorithms to identify vulnerabilities in the system, and the implementation of encryption schemes as a response or as a preventive measure for all types of CPS. Our work on the other hand presents an overview of all the model-based techniques used for evaluating ACPPS's trustworthiness.

A recent decade-wide literature review published provides an overview of security aspects in **Industrial CPS** (**I-CPS**) [3]. The study, however, focuses only on discussing security-related properties. It highlights the identification methods, such as attack detection and finally presents a summary of the proposed mitigation measures.

A survey on security in CPS and Industry 4.0 published by Yaccoub et al. [85] provides an overview of vulnerabilities, threats, attacks, and failures in IIoT, proposed measures along with their limitations. According to this study, the threats can be categorized as cyber and physical threats. For the discussed threats, the survey summarise the non-model-based tools available to evaluate systems for such threats.

Another article published in the year 2019 [4] discusses the formation of ACPPS followed by the types of cyber-security threats that exist, making the system vulnerable. The Source of the attack being internal or external, the use of **software-defined networks** (**SDN**) and **network function virtualization** (**NFV**) can assist in automatic incident response by rapidly detecting and replacing the failed component with its virtual implementation.

Geismann et al., in the published systematic literature review, summarizes aspects of model-driven engineering in CPS. The study presents an overview of 7 specific model-based tools that allow for modeling and evaluation of the system at an early stage of software development. Although this particular study is relevant to our work, it is focused on CPS as a whole and not ACPPS, and it discusses the analysis of only security-related properties [46].
A recent study published by Gaiardelli [44] discusses the adaptation of Model-based Software Engineering to developed software related to Industry 5.0. It provides an overview of state-of-the-art tools proposed to model AI-integrated CPS using SysML modeling language. The tools presented allow not only the construction but also the evaluation of the system. However, the study does not specify the type of properties the system is evaluated for. In the meantime, our work provides an overview of all possible modeling languages used to evaluate trustworthiness attributes in ACPPS. In general, all of these studies focus only on one particular aspect of trustworthiness, i.e., security, whereas our study further extends this type of research onto other aspects of trustworthiness as well. Our work also provides an overview of the model-based techniques and tools used for evaluation purposes. The literature presented in this section lacks a discussion on the model-based techniques proposed over the years to evaluate ACPPS's trustworthiness.

7 CONCLUSION

This systematic mapping study presents a structured understanding of the state-of-the-art model-based techniques used for evaluating trustworthiness in ACPPS. During this study, we identified, classified, and analyzed articles published till June 2022 using precise data extraction, analysis, and synthesis processes. From the initial set of 959 peer-reviewed publications, we reached a final set of 44 primary studies. The main findings of this mapping study are as follows:

- The articles focusing on the model-based evaluation of trustworthiness in ACPPS started to be published in 2016, from where it gradually started to gain popularity, reaching a peak in 2020 with a total of 10 articles published. Most of these articles published were either conference papers or journal articles.
- Since the domain is relatively new, not all aspects of trustworthiness defined in existing taxonomies have been evaluated, especially using model-based techniques in ACPPS. However, some of the most reported aspects of trustworthiness being evaluated consisted of Safety and Security as a whole.
- Looking into the model formalism, Petri Nets, SysML, and UML were among the most reported modeling languages used to construct models of the overall system or a subsystem. Some articles also used programming languages such as Python, FORTRAN, and G-Code to construct meta-models to be used to construct a system's model.
- Articles reported using tools that were either model-based or other detection and development tools. Petri Net tools, MATLAB, and CAD were among the most reported model-based tools used for evaluation purposes, while Integrated Development Tools were among the most reported programming tools used to assist in the evaluation process of ACPPS.
- Not all the articles presented the working of their proposed model-based evaluation techniques. However, those that did mainly used simulations of the entire system, sub-systems, or specific scenarios as a case study.

8 FUTURE WORK

Based on our findings, we believe that the domain is relatively new, with several emerging and promising approaches based on multi-paradigm modeling, DSL, and M2M transformations. Engineers need to tackle many challenges to leverage the potential of model-based evaluation in

ACPPS, considering their complexity, heterogeneity, and limited predictability/transparency. That is mainly due to the usage of AI, and specifically ML, in order to manage autonomy against uncertainties as well as system learning and evolution capabilities to dynamically adapt to changes in the operating environments.

One future research opportunity could be the investigation into the limitations of the proposed model-based evaluation techniques presented in this study in the domain of ACPPS. The aim here will be to observe the extent to which the proposed techniques can evaluate the trustworthiness-related attributes in ACPPS, more specifically in the AI component of ACPPS.

Since not all trustworthiness-related attributes defined as per existing taxonomies are reported to have been evaluated using model-based techniques, an investigation into the alternative solutions used for their evaluation and the reasons behind them could be conducted to help better understand the shortcomings of the techniques highlighted in this study.

Another possible research opportunity could be to study the possibility of applying model-based evaluation techniques mainly defined for a generic CPS to analyze trustworthiness attributes implemented in ACPPS.

REFERENCES

[1] 2020. The CIA Triad — Confidentiality, Integrity, and Availability Explained. Retrieved February 1st, 2020 from https://www.freecodecamp.org/news/the-cia-triad-confidentiality-integrity-and-availability-explained/

[2] Mowafak H. Abdu-Hussin. 2016. An S4PR class petri net supervisor for manufacturing system. *International Journal of Simulation: Systems, Science and Technology* 17, 33 (2016), 31.1–31.11. DOI: https://doi.org/10.5013/IJSSST.a.17.33.31

[3] Neha Agrawal and Rohit Kumar. 2022. Security perspective analysis of industrial cyber physical systems (I-CPS): A decade-wide survey. 130, 0019–0578 (2022), 10–24. https://doi.org/10.1016/j.isatra.2022.03.018

[4] V. Alcácer and V. Cruz-Machado. 2019. Scanning the Industry 4.0: A Literature Review on Technologies for Manufacturing Systems. *Engineering Science and Technology, an International Journal* 22, 3 (2019), 899–919 pages. DOI: https://doi.org/10.1016/j.jestch.2019.01.006

[5] Temesgen Seyoum Alemayehu, Jai Hoon Kim, and We Duke Cho. 2021. Optimal replacement model for the physical component of safety critical smart-world CPSs. *Journal of Ambient Intelligence and Humanized Computing* 13, 1868–5145 (2021), 1–12. DOI: https://doi.org/10.1007/s12652-021-03137-5

[6] Nauman Bin Ali and Kai Petersen. 2014. Evaluating strategies for study selection in systematic literature studies. In *Procs of ESEM*. ACM.

[7] Bhagyashree Ambore and L. Suresh. 2019. Novel model for boosting security strength and energy efficiency in internet-of-things using multi-staged game. *International Journal of Electrical and Computer Engineering* 9, 5 (2019), 4326–4335. DOI: https://doi.org/10.11591/ijece.v9i5.pp4326-4335

[8] Mihai Andronie, George Lăzăroiu, Mariana Iatagan, Cristian Uță, Roxana Ștefănescu, and Mădălina Cocoșatu. 2021. Artificial intelligence-based decision-making algorithms, internet of things sensing networks, and deep learning-assisted smart process management in cyber-physical production systems. *Electronics* 10, 2079–9292 (2021), 2497.

[9] Michael Backes, Jannik Dreier, Steve Kremer, and Robert Kunnemann. 2017. Security viewpoint in a reference architecture model for cyber-physical production systems. *2nd IEEE European Symposium on Security and Privacy, EuroS and P* (2017), 76–91. https://doi.org/10.1109/EuroSP.2017.12

[10] Torsten Bandyszak, Marian Daun, Bastian Tenbergen, Patrick Kuhs, Stefanie Wolf, and Thorsten Weyer. 2020. Orthogonal uncertainty modeling in the engineering of cyber-physical systems. *IEEE Transactions on Automation Science and Engineering* 17, 3 (2020), 1250–1265. DOI: https://doi.org/10.1109/TASE.2020.2980726

[11] Torsten Bandyszak, Marian Daun, Bastian Tenbergen, and Thorsten Weyer. 2018. Model-based documentation of context uncertainty for cyber-physical systems (an approach and application to an industry automation case example). *IEEE 14th International Conference on Automation Science and Engineering (CASE)*. https://doi.org/10.0/Linux-x86_64

[12] Ankica Barišić, Ivan Ruchkin, Dušan Savić, Mustafa Abshir Mohamed, Rima Al-Ali, Letitia W. Li, Hana Mkaouar, Raheleh Eslampanah, Moharram Challenger, Dominique Blouin, Oksana Nikiforova, and Antonio Cicchetti. 2022. Multi-paradigm modeling for cyber–physical systems: A systematic mapping review. *Journal of Systems and Software* 183, 0164–1212 (2022), 111081. https://doi.org/10.1016/j.jss.2021.111081

[13] Victor R. Basili, Gianluigi Caldiera, and H. Dieter Rombach. 1994. The goal question metric approach. In *Encyclopedia of Software Engineering*. Vol. 2. Wiley, 528–532.

[14] Sara Moghadaszadeh Bazaz, Mika Lohtander, and Juha Varis. 2019. 5-dimensional definition for a manufacturing digital twin. *29th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM'19)* 38, 2351–9789 (2019), 1705–1712. https://doi.org/10.1016/j.promfg.2020.01.107

[15] Christian Berger, Philipp Eichhammer, Hans P. Reiser, Jörg Domaschka, Franz J. Hauck, and Gerhard Habiger. 2022. A survey on resilience in the IoT: Taxonomy, classification, and discussion of resilience mechanisms. *Comput. Surveys* 54, 7 (2022), 1–39. https://doi.org/10.1145/3462513

[16] Stephan Berger, Bofenreuther Maximilian, Häckel Björn, and Oliver Niesel. 2019. Modelling availability risks of IT threats in smart factory networks-a modular petri net approach. *27th European Conference on Information Systems (ECIS)*. https://aisel.aisnet.org/ecis2019_rp

[17] Stephan Berger, Christopher van Dun, and Björn Häckel. 2022. IT availability risks in smart factory networks - analyzing the effects of IT threats on production processes using petri nets. *Information Systems Frontiers* 24, 1572–9419 (2022), 1–20. https://doi.org/10.1007/s10796-02210243-y

[18] Elias Bou-Harb, Evangelos I. Kaisar, and Mark Austin. 2017. On the impact of empirical attack models targeting marine transportation. *5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*.

[19] Alejandro Bracho, Can Saygin, Hungda Wan, Yooneun Lee, and Alireza Zarreh. 2018. A simulation-based platform for assessing the impact of cyber-threats on smart manufacturing systems. *46th SME North American Manufacturing Research Conference (NAMRC 46)*, Vol. 26, 1116–1127. https://doi.org/10.1016/j.promfg.2018.07.148

[20] Pearl Brereton, Barbara A. Kitchenham, David Budgen, Mark Turner, and Mohamed Khalil. 2007. Lessons from applying the systematic literature review process within the software engineering domain. *Journal of Systems and Software* 80, 0164–1212 (2007), 571–583.

[21] Alessio Bucaioni, Francesco Flammini, and Mats Ahlskog. 2020. Towards model-based performability evaluation of production systems. *25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. 1946–0759.

[22] Alessio Bucaioni, Francesco Flammini, and Mats Ahlskog. 2020. Towards model-based performability evaluation of production systems. In *2020 25th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*, Vol. 1. IEEE, 1085–1088.

[23] John H. Castellanos, Martín Ochoa, and Jianying Zhou. 2018. Finding dependencies between cyber-physical domains for security testing of industrial control systems. *Annual Computer Security Applications Conference*, 582–594. https://doi.org/10.1145/3274694.3274745

[24] Vladana Čelebić and Alessio Bucaioni. 2023. A systematic mapping study on the role of software engineering in enabling society 5.0. In *2023 IEEE International Smart Cities Conference (ISC2)*. IEEE, 1–8.

[25] Kathy Charmaz and Linda Liska Belgrave. 2007. Grounded theory. *The Blackwell Encyclopedia of Sociology* 11 (2007), 2023–2027.

[26] Dejiu Chen, Dmitri Valeri Panfilenko, Mahmood R. Khabbazi, and Daniel Sonntag. 2016. A model-based approach to qualified process automation for anomaly detection and treatment. *IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'16-November)*. 1946–0759. https://doi.org/10.1109/ETFA.2016.7733731

[27] Lianping Chen, Muhammad Ali Babar, and Bashar Nuseibeh. 2012. Characterizing architecturally significant requirements. *IEEE Software* 30, 1937–4194 (2012), 38–45.

[28] Sujit Rokka Chhetri, Anthony Bahadir Lopez, Jiang Wan, and Abdullah Al Faruque. 2019. GAN-Sec: Generative adversarial network modeling for the security analysis of cyber-physical production systems. *Design, Automation & Test in Europe Conference & Exhibition (DATE)*.

[29] European Commision. 2022. What is Industry 5.0? Retrieved from https://research-and-innovation.ec.europa.eu/research-area/industry/industry-50_en1/4. Access Date: 2022.

[30] European Commission, Directorate-General for Communications Networks, Content, and Technology. 2022. HIGH-LEVEL EXPERT GROUP ON ARTIFICIAL INTELLIGENCE SET UP BY THE EUROPEAN COMMISSION ETHICS GUIDELINES FOR TRUSTWORTHY AI. https://ec.europa.eu/digital- Access Date: June 7th, 2022.

[31] Daniela S. Cruzes and Tore Dyba. 2011. Recommended steps for thematic synthesis in software engineering. In *Procs of ESEM*. IEEE, 275–284.

[32] Marian Daun, Jennifer Brings, Thorsten Weyer, and Bastian Tenbergen. 2016. Fostering concurrent engineering of cyber-physical systems: A proposal for an ontological context framework. *3rd International Workshop on Emerging Ideas and Trends in Engineering of Cyber-Physical Systems (EITEC'16)*. 5–10. https://doi.org/10.1109/EITEC.2016.7503689

[33] Zach DeSmit, Ahmad E. Elhabashy, Lee J. Wells, and Jaime A. Camelio. 2016. Cyber-physical vulnerability assessment in manufacturing systems. *44th Proceedings of the North American Manufacturing Research Institution of SM* 5, 2351–9789 (2016), 1060–1074. https://doi.org/10.1016/j.promfg.2016.08.075

[34] Zach DeSmit, Ahmad E. Elhabashy, Lee J. Wells, and Jaime A. Camelio. 2017. An approach to cyber-physical vulnerability assessment for intelligent manufacturing systems. *Journal of Manufacturing Systems* 43, 0278–6125 (2017), 339–351. https://doi.org/10.1016/j.jmsy.2017.03.004

[35] Zach DeSmit, Aditya U. Kulkarni, and Christian Wernz. 2018. Enhancing cyber-physical security in manufacturing through game-theoretic analysis. *Cyber-Physical Systems* 4, 4 (2018), 232–259. DOI: https://doi.org/10.1080/23335777.2018.1537302

[36] Romina Eramo and Alessio Bucaioni. 2013. Understanding bidirectional transformations with TGGs and JTL. *Electronic Communications of the EASST* 57, 1863–2122 (2013), 1–20.

[37] Tagir Fabarisov, Nafisa Yusupova, Kai Ding, Andrey Morozov, and Klaus Janschek. 2020. Model-based stochastic error propagation analysis for cyber-physical systems. *Acta Polytechnica Hungarica* 17, 8 (2020), 15–28. DOI: https://doi.org/10.12700/APH.17.8.2020.8.2

[38] Enxhi Ferko, Alessio Bucaioni, and Moris Behnam. 2022. Architecting digital twins. *IEEE Access* 10, 2169–3536 (2022), 50335–50350. DOI: https://doi.org/10.1109/ACCESS.2022.3172964

[39] Enxhi Ferko, Alessio Bucaioni, Patrizio Pelliccione, and Moris Behnam. 2023. Standardisation in digital twin architectures in manufacturing. In *2023 IEEE 20th International Conference on Software Architecture (ICSA)*. 70–81. https://doi.org/10.1109/ICSA56044.2023.00015

[40] Francesco Flammini, Cristina Alcaraz, Emanuele Bellin, Stefano Marrone, Javier Lopez, and Andrea Bondavalli. 2022. Towards trustworthy autonomous systems: A survey of taxonomies and future perspectives. *IEEE Transactions on Emerging Topics in Computing* 2168–6750 (2022), 1–13.

[41] Roberto Franzosi. 2010. *Quantitative Narrative Analysis*. Sage.

[42] Josh Fruhlinger. 2020. The CIA triad: Definition, components and examples. Retrieved February 2020 from https://www.csoonline.com/article/568917/the-cia-triad-definition-components-and-examples.html

[43] Luca Fumagalli, Marco Macchi, Cristian Colace, Maurizio Rondi, and Alessandro Alfieri. 2016. A smart maintenance tool for a safe electric arc furnace. *Part of Special Issue: 12th IFAC Workshop on Intelligent Manufacturing Systems IMS* 49, 31 (2016), 19–24. DOI: https://doi.org/10.1016/j.ifacol.2016.12.155

[44] Sebastiano Gaiardelli, Stefano Spellini, Michele Lora, and Franco Fummi. 2021. Modeling in industry 5.0: What is there and what is missing: Special session 1: Languages for industry 5.0. *Forum on Specification and Design Languages* 2021-September. 1636–9874, 01–08. https://doi.org/10.1109/FDL53530.2021.9568371

[45] Vahid Garousi, Michael Felderer, and Mika V. Mäntylä. 2019. Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology* 106, 0950–5849 (2019), 101–121.

[46] Johannes Geismann and Eric Bodden. 2020. A systematic literature review of model-driven security engineering for cyber–physical systems. *Journal of Systems and Software* 169, 0164–1212 (2020), 110–697. DOI: https://doi.org/10.1016/j.jss.2020.110697

[47] Jairo Giraldo, Esha Sarkar, Alvaro A. Cardenas, Michail Maniatakos, and Murat Kantarcioglu. 2017. Security and privacy in cyber-physical systems: A survey of surveys. *IEEE Design and Test* 34, 4 (2017), 7–17. DOI: https://doi.org/10.1109/MDAT.2017.2709310

[48] Carlos A. Gonzalez, Mojtaba Varmazyar, Shiva Nejati, Lionel C. Briand, and Yago Isasi. 2020. Enabling model testing of cyber-physical systems. *21th ACM/IEEE International Conference on Model Driven Engineering Languages and Systems* 2657, 1613–0073 (2020), 1–9. https://doi.org/10.1145/nnnnnnn.nnnnnnn

[49] Trisha Greenhalgh and Richard Peacock. 2005. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: Audit of primary sources. *BMJ* 331, 7524 (2005), 1064–1065.

[50] Hao Hu, Xiaoliang Jia, Kuo Liu, and Bingyang Sun. 2021. Self-adaptive traffic control model with behavior trees and reinforcement learning for AGV in industry 4.0. *IEEE Transactions on Industrial Informatics* 17, 1941–0050 (2021), 7968–7979. https://doi.org/10.1109/TII.2021.3059676

[51] Mian Ahmad Jan, Fazlullah Khan, Rahim Khan, Spyridon Mastorakis, Varun G. Menon, Mamoun Alazab, and Paul Watters. 2021. Lightweight mutual authentication and privacy-preservation scheme for intelligent wearable devices in industrial-CPS. *IEEE Transactions on Industrial Informatics* 17, 8 (2021), 5829–5839. DOI: https://doi.org/10.1109/TII.2020.3043802

[52] Sergej Japs, Harald Anacker, and Roman Dumitrescu. 2021. SAVE: Security & safety by model-based systems engineering on the example of automotive industry. *31st CIRP Design Conference* 100, 2212–8271 (2021), 187–192. https://doi.org/10.1016/j.procir.2021.05.053

[53] Mohammad Jbair, Bilal Ahmad, Carsten Maple, and Robert Harrison. 2022. Threat modelling for industrial cyber physical systems in the era of smart manufacturing. *Computers in Industry* 137, 0166–3615 (2022), 103–611. DOI: https://doi.org/10.1016/j.compind.2022.103611

[54] Yuning Jiang, Yacine Atif, Jianguo Ding, and Wei Wang. 2019. A semantic framework with humans in the loop for vulnerability-assessment in cyber-physical production systems. *14th International Conference on Risks and Security of Internet and Systems 12026 LNCS*, 1611–3349, 128–143. https://doi.org/10.1007/978-3-030-41568-6_9

[55] Georgios Kavallieratos and Sokratis Katsikas. 2020. Attack path analysis for cyber physical systems. 19–33. Retrieved from http://www.springer.com/series/7410

[56] Georgios Kavallieratos and Sokratis Katsikas. 2020. Attack path analysis for cyber physical systems. In *International Workshop on the Security of Industrial Control Systems and Cyber-Physical Systems, International Workshop on Security and Privacy Requirements Engineering, International Workshop on Attacks and Defenses for Internet-of-Things*. Springer, 19–33.

[57] Matthias Kern, Emre Taspolatoglu, Fabian Scheytt, Thomas Glock, Bo Liu, Victor Pazmino Betancourt, Juergen Becker, and Eric Sax. 2020. An architecture-based modeling approach using data flows for zone concepts in industry 4.0. In *Proceedings of the ISSE 2020 - 6th IEEE International Symposium on Systems Engineering*. DOI: https://doi.org/10.1109/ISSE49799.2020.9272013

[58] Ateeq Khan and Klaus Turowski. 2016. A perspective on industry 4.0: From challenges to opportunities in production systems. *IoTBD 2016 - Proceedings of the International Conference on Internet of Things and Big Data 978–989–758–183–0*, 441–448. https://doi.org/10.5220/0005929704410448

[59] Izhar Ahmed Khan, Marwa Keshk, Dechang Pi, Nasrullah Khan, Yasir Hussain, and Hatem Soliman. 2022. Enhancing IIoT networks protection: A robust security model for attack detection in internet industrial control systems. *Ad Hoc Networks* 134, 1570–8705 (2022), 102930. DOI: https://doi.org/10.1016/j.adhoc.2022.102930

[60] Barbara Kitchenham and Pearl Brereton. 2013. A systematic review of systematic review process research in software engineering. *Information and Software Technology* 55, 0950–5849 (2013), 2049–2075.

[61] Barbara A. Kitchenham and Stuart Charters. 2007. *Guidelines for Performing Systematic Literature Reviews in Software Engineering*. Technical Report EBSE-2007-01. Keele University and University of Durham.

[62] Peter Gorm Larsen, John Fitzgerald, Jim Woodcock, Peter Fritzson, Jörg Brauer, Christian Kleijn, Thierry Lecomte, Markus Pfeil, Ole Green, Stylianos Basagiannis, et al. 2016. Integrated tool chain for model-based design of Cyber-Physical Systems: The INTO-CPS project. In *2nd International Workshop on Modelling, Analysis, and Control of Complex CPS (CPS Data)*. IEEE, 1–6.

[63] Daniel Lichte and Kai-Dietrich Wolf. 2018. Use case-based consideration of safety and security in cyber physical production systems applied to a collaborative robot system. In *Safety and Reliability–Safe Societies in a Changing World*, 1395–1401. Retrieved June 2018 from https://www.researchgate.net/publication/325654823

[64] Professor Messer. 2020. Confidentiality, integrity, availability, and safety - comptia security+ sy0-401: 2.9. Retrieved from https://www.professormesser.com/security-plus/sy0-401/confidentiality-integrity-availability-and-safety/

[65] Nazila Gol Mohammadi, Mohamed Bishr, Andreas Metzger, Thorsten Weyer, Klaus Pohl, Sachar Paulus, Holger Könnecke, and Sandro Hartenstein. 2014. Trustworthiness attributes and metrics for engineering trusted internet-based software systems. In *International Conference on Cloud Computing and Services Science*. 19–35. http://www.springer.com/series/7899

[66] László Monostori. 2014. Cyber-physical production systems: Roots, expectations and R&D challenges. *Procedia Cirp* 17 (2014), 9–13.

[67] Nour Moustafa, Erwin Adi, Benjamin Turnbull, and Jiankun Hu. 2018. A new threat intelligence scheme for safeguarding industry 4.0 systems. *IEEE Access* 6, 2169–3536 (2018), 32910–32924. DOI: https://doi.org/10.1109/ACCESS.2018.2844794

[68] Lin Mu, Enjin Zhao, Yuewei Wang, and Albert Y. Zomaya. 2020. Buoy sensor cyberattack detection in offshore petroleum cyber-physical systems. *IEEE Transactions on Services Computing* 13, 4 (2020), 653–662. DOI: https://doi.org/10.1109/TSC.2020.2964548

[69] Hamed Orojloo and Mohammad Abdollahi Azgomi. 2017. A game-theoretic approach to model and quantify the security of cyber-physical systems. *Computers in Industry* 88, 0166–3615 (2017), 44–57. DOI: https://doi.org/10.1016/j.compind.2017.03.007

[70] Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. 2008. Systematic mapping studies in software engineering. In *Proceedings of the EASE*. 68–77.

[71] Davy Preuveneers, Wouter Joosen, and Elisabeth Ilie-Zudor. 2018. Robust digital twin compositions for industry 4.0 smart manufacturing systems. In *IEEE International Enterprise Distributed Object Computing Workshop, (EDOCW)*, Vol. 2018-October. Institute of Electrical and Electronics Engineers Inc., 69–78. https://doi.org/10.1109/EDOCW.2018.00021

[72] CDI Products. 2021. How Cobots Are Powering Smart Manufacturing.

[73] Lokesh Kumar Rathore and Neelabh Sao. 2015. An integrated model based test case prioritization using UML sequence and activity diagram. *International Journal of Research in Computer Applications and Robotics* 3, 2320–7345 (2015), 31–41. Access Date: December 2015.

[74] Vinod Saratchandran. 2022. Cobots and The Future of Manufacturing: A Quick Glimpse! Retrieved from https://www.fingent.com/blog/cobots-and-the-future-of-manufacturing-a-quick-glimpse/

[75] Nathalie A. Smuha. 2019. The EU approach to ethics guidelines for trustworthy artificial intelligence: A continuous journey towards an appropriate governance framework for AI. *A Journal of Information Law and Technology* (2019), 97–106. Retrieved from https://www.mmcventures.com/wp-content/uploads/2019/0

[76] Daniel Stock, Daniel Schel, and Thomas Bauernhansl. 2019. Cyber-physical production system self-description-based data access layer. In *24th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA)*. 168–175. https://doi.org/10.1109/ETFA.2019.8869486

[77] Sebastian Thiede. 2021. Cyber-physical production systems (CPPS): introduction. 24 pages.

[78] Tagline Treichel, Pablo Oliveira Antonino, Filipe Silva Santos, and Leonardo Silva Rosa. 2021. Simulation-as-a-service: A simulation platform for cyber-physical systems. *2021 IEEE 18th International Conference on Software Architecture Companion (ICSA-C'21)*. 155–161. https://doi.org/10.1109/ICSA-C52384.2021.00038

[79] Edson H. Watanabe, Robson M. da Silva, Fabricio Junqueira, Diolino J. dos Santos Filho, and Paulo E. Miyagi. 2016. An emerging industrial business model considering sustainability evaluation and using cyber physical system technology and modelling techniques. *Cyber-Physical and Human-Systems (CPHS'16)* 49, 32 (2016), 135–140. DOI: https://doi.org/10.1016/j.ifacol.2016.12.203

[80] Alejandro White, Ali Karimoddini, and Mohammad Karimadini. 2020. Resilient fault diagnosis under imperfect observations-a need for industry 4.0 era. *IEEE/CAA Journal of Automatica Sinica* 7, 5 (2020), 1279–1288. DOI: https://doi.org/10.1109/JAS.2020.1003333

[81] Claes Wohlin. 2014. Guidelines for snowballing in systematic literature studies and a replication in software engineering. In *Procs of EASE*. ACM, 10 pages. https://doi.org/10.1145/2601248.2601268

[82] C. Wohlin, P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, and A. Wesslén. 2012. *Experimentation in Software Engineering*. Springer.

[83] Zhitao Wu, Xiaoming Yang, Ping Chen, Zongshun Qu, and Jun Lin. 2021. Multi-scale software network model for software safety of the intended functionality. *IEEE International Symposium on Software Reliability Engineering Workshops (ISSREW'21)*. 250–255. https://doi.org/10.1109/ISSREW53611.2021.00071

[84] Xin Xin, Sye Loong Keoh, Michele Sevegnani, and Martin Saerbeck. 2020. Dynamic probabilistic model checking for sensor validation in industry 4.0 applications. *IEEE International Conference on Smart Internet of Things, SmartIoT 2020 978–1–7281–6514–1*, 43–50. https://doi.org/10.1109/SmartIoT49966.2020.00016

[85] Jean Paul A. Yaacoub, Ola Salman, Hassan N. Noura, Nesrine Kaaniche, Ali Chehab, and Mohamad Malli. 2020. Cyber-physical systems security: Limitations, issues and future trends. *Microprocessors and Microsystems* 77, 0141–9331 (2020). DOI: https://doi.org/10.1016/j.micpro.2020.103201

[86] Zhenhua Yu, Lijun Zhou, Zhiqiang Ma, and Mohammed A. El-Meligy. 2017. Trustworthiness modeling and analysis of cyber-physical manufacturing systems. *IEEE Access* 5, 2169–3536 (2017), 26076–26085. DOI: https://doi.org/10.1109/ACCESS.2017.2777438

[87] Alireza Zarreh, Yooneun Lee, Rafid Al Janahi, Hung Da Wan, and Can Saygin. 2020. Cyber-physical security evaluation in manufacturing systems with a bayesian game model. In *Proceedings of the 30th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM2021)*. 51, 2351–9789 (2020), 1158–1165. DOI: https://doi.org/10.1016/j.promfg.2020.10.163

[88] Alireza Zarreh, Can Saygin, HungDa Wan, Yooneun Lee, Alejandro Bracho, et al. 2018. Cybersecurity analysis of smart manufacturing system using game theory approach and quantal response equilibrium. *Procedia Manufacturing* 17 (2018), 1001–1008.

Received 15 October 2022; revised 10 October 2023; accepted 21 December 2023