An integrated safety and security analysis for cyber-physical harm scenarios

Carreras Guzman, Nelson H.; Kozine, Igor; Lundteigen, Mary Ann

Published in:
Safety Science

Link to article, DOI:
10.1016/j.ssci.2021.105458

Publication date:
2021

Document Version
Publisher's PDF, also known as Version of record

Link back to DTU Orbit

Citation (APA):
Carreras Guzman, N. H., Kozine, I., & Lundteigen, M. A. (2021). An integrated safety and security analysis for cyber-physical harm scenarios. *Safety Science*, *144*, [105458]. https://doi.org/10.1016/j.ssci.2021.105458

Safety Science 144 (2021) 105458

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

An integrated safety and security analysis for cyber-physical harm scenarios

Nelson H. Carreras Guzman [a,b,*], Igor Kozine [a], Mary Ann Lundteigen [c]

[a] Engineering Systems Design Group, Technical University of Denmark (DTU), Kgs. Lyngby 2800, Denmark
[b] Department of Mechanical and Industrial Engineering, Norwegian University of Science and Technology (NTNU), Trondheim 7034, Norway
[c] Department of Engineering Cybernetics, Norwegian University of Science and Technology (NTNU), Trondheim 7034, Norway

ARTICLE INFO

Keywords:
Cyber-Physical Systems (CPSs)
Autonomous Systems
Cyber-Physical Harm Analysis for Safety and Security (CyPHASS)
Bowtie Method

ABSTRACT

Increasing digitalization and autonomous solutions in physical systems promise to enhance their performance, cost-efficiency and reliability. However, the integration of novel information technologies with safety-related systems also brings new vulnerabilities and risks that challenge the traditional field of safety analysis. Particularly, cyber security threats are becoming key factors in complex accident scenarios in cyber-physical systems (CPSs), where unintentional errors and design flaws overlap with cyber security vulnerabilities that could lead to harm to humans and assets. This overlap between safety and security analysis is still a loosely defined domain without established theories and methods, leading to complications during the risk analysis of CPSs. In this paper, we first describe how the domain of safety science increasingly overlaps with security analysis. Subsequently, based on this overlapping, we illustrate and complement an integrated method for the identification of harm scenarios in CPSs. This method, coined Uncontrolled Flows of Information and Energy (UFoI-E), offers a distinct theoretical foundation rooted in accident causation models and a framework to design diagrammatic representations of CPSs during the analysis. After summarizing these features of the UFoI-E method, we present our original contribution to the method, which is a new practical toolkit for risk identification composed of an ontology of harm scenarios and a database of checklists built from lessons learned analysis and expert knowledge. Finally, we demonstrate an application of the method in an illustrative case and show representative fields for future work.

1. Introduction

Significant incentives are leading to the integration of novel digital technologies in the architectures of physical systems, enabling better performance, cost efficiency, energy efficiency, and other attractive benefits. Scholars have associated these technological developments with the 4th Industrial revolution of cyber-physical systems (CPSs), transforming the operation of physical systems with increased automation, connectivity and smartness (Jazdi, 2014; Monostori, 2014). Hellen Gill was the first to propose the term CPS at the National Science Foundation of the United States in 2006, where she highlighted some of the challenges in the operations of engineered systems that are tightly monitored and controlled by a computational core (Gill, 2008; Lee, 2006). Since then, the concept has been discussed and interpreted a bit differently among researchers and in industry (Park et al., 2012).

Some researchers have proposed approaches to design, operate, monitor and protect these innovative systems in suitable ways with a focus on embedded system design (Lee, 2008; Marwedel, 2011). Nevertheless, these embedded systems are not isolated and purely autonomous components; instead, they are part of larger CPS architectures that interact with human operators and with the surrounding physical environment (Rajkumar et al., 2017; Rajkumar et al., 2010).

In this paper, we have adopted the definition that CPSs are "engineered systems that integrate information technologies, real-time control subsystems, physical components and human operators to influence physical processes by means of cooperative and (semi)automated control functions" (Carreras Guzman et al., 2020). In this view, the key features of CPSs are the intersection between:

(1) "Real-time feedback control of physical processes through sensors and actuators
(2) Cooperative control among networked subsystems, and
(3) A threshold of automation level where computers close the feedback control loops in (semi)automated tasks, possibly allowing human control in certain cases." (Carreras Guzman et al., 2020)

* Corresponding author at: Akademivej, Building 358, Kgs. Lyngby 2800, Denmark.
E-mail address: nelca@dtu.dk (N.H. Carreras Guzman).

https://doi.org/10.1016/j.ssci.2021.105458
Received 1 April 2020; Received in revised form 4 June 2021; Accepted 21 August 2021
Available online 30 August 2021

Many CPSs are safety-related, that is, they interact with the physical world and could potentially lead to physical harm scenarios killing or injuring humans, damaging assets or affecting natural ecosystems. Examples include industrial plants such as petro-chemical refinement plants, electric power stations and manufacturing plants (Ge et al., 2017; Khaitan and McCalley, 2015; Monostori et al., 2016). Furthermore, mainly all transportation sectors are also evolving with digitalization, showing the rise of autonomous cars and remotely controlled vessels as well as new developments in (semi)autonomous aircrafts, trains and mobile machinery (Fridman, 2019; Hagen et al., 2018; Lee and Seshia, 2017). Other safety-related CPSs include smart medical devices (Rajkumar et al., 2017), robotics (Michniewicz and Reinhart, 2014), and military weapon systems (Protti and Barzan, 2007).

The integration of digital technologies tends to increase the complexity of the system, and, in many cases, it also introduces new security vulnerabilities that system designers could easily overlook. Indeed, emerging risks in safety-related CPSs include both unintentional and intentional risk sources (Friedberg et al., 2017; Zio, 2018).

As emerging security risks, this paper refers to *cyber-physical attacks* as intentional cyber threats that aim at disrupting physical processes and potentially lead to human injuries, asset damages and/or impacts to natural ecosystems. In other words, cyber-physical attacks are the subset of cyber-attacks intentionally willing to provoke physical harm (He and Yan, 2016). In the Danish Defence Intelligence Service, these cyber-attacks are also known as *destructive cyber-attacks* (Centre for Cyber Security, 2019). Recent examples include real world cyber-physical attacks against process plants, as well as research experiments proving new attack surfaces in vehicles, medical devices, among other systems (Humayed et al., 2017; Weiss, 2010).

In response to the rising cases of cyber-physical attacks, researchers have recently proposed methods to integrate safety and security analysis. A review of these existing methods is provided in Section 2.1. Grounded on this review, we conclude that there are two main limitations in the existing methods. First, some methods tend to be suitable only for early system lifecycle phases or for simple system architectures. And second, some methods do not integrate the relationships between the information systems and the specific energies controlled by the CPSs in the physical world. To overcome these limitations, we have developed the UFoI-E method, and this paper finalizes its presentation introducing the new constituent, the CyPHASS scenario builder, and providing a consistent view on the method that unities all its constituents.

This paper is organized as follows. Section 2 describes the challenges and explains the overlaps between the fields of safety analysis and cyber-physical security. Section 3 presents an integrated method for safety and security analysis of CPSs. Section 4 explains our main contribution to this integrated method, which is an ontology of harm scenarios and a database of checklists that analysts can use as a practical tool for systematic risk identification. Section 5 demonstrates an application of this integrated method in an illustrative case study in the nuclear safety domain and mentions other domains of ongoing research. Finally, Section 6 concludes.

2. Overlaps of safety and security in CPSs

According to Aven (2014), we can understand safety science as the "knowledge about safety related issues, and the development of concepts, theories, principles and methods to understand, assess, communicate and manage (in a broad sense) safety". In this sense, the literature associates unintentional or random risk sources (hazards) to the domain of safety, and intentional and malicious risk sources (threats) to the domain of security (Amundrud et al., 2017; Paul et al., 2016; Pietre-Cambacedes and Chaudet, 2010).

Furthermore, there is a persistent distinction in the type of undesired outcomes that safety and security analysts tend to address. Especially in the domain of cyber security, the focus has traditionally been to prevent the violation of security properties. These security properties are usually summarized as the data confidentiality, integrity, and availability triad (CIA) and can be expanded with the inclusion of properties such as authenticity, nonrepudiation, among others (Dzung et al., 2005; IEC/TS 62443-1-1, 2009; ISO/IEC 27000, 2017; Stouffer et al., 2015). However, the emerging cyber-physical attacks challenge this limited scope of cyber security and expand the type of undesired outcomes to include physical harm scenarios associated to safety. Therefore, Aven's (2014) remark that the definition of safety could optionally include intentional incidents is not only an option, but a critical requirement in the domain of CPSs.

Hence, we argue that the domain of safety science should not only comprise hazards as risk sources that could lead to physical harm scenarios. Instead, there is a need to include the subdomain described by Stephane Paul as "security for safety" (Paul, 2015; Paul et al., 2016), incorporating the subset of security threats leading to physical harm scenarios. The security for safety domain should include both physical attacks (traditionally associated to sabotage and terrorism) and cyber-physical attacks. In Fig. 1, we illustrate this integration of security for safety in the domain of safety science.

The integration will facilitate identification of the propagation routes linking security breaches and hazards. In this way, it will be possible to introduce measures to interrupt the propagation, which in turn will contribute to making CPSs safer. Expanding on a similar point, Max Tegmark - a leading expert in smart systems and artificial intelligence – writes:

> "Now that our machines are getting smart enough to have some information about what they are doing, it is time for us to teach them limits. Any engineer designing a machine needs to ask if there are things that it can but should not do, and consider whether there is practical way of making it impossible for a malicious or clumsy user to cause harm" (Tegmark, 2018)

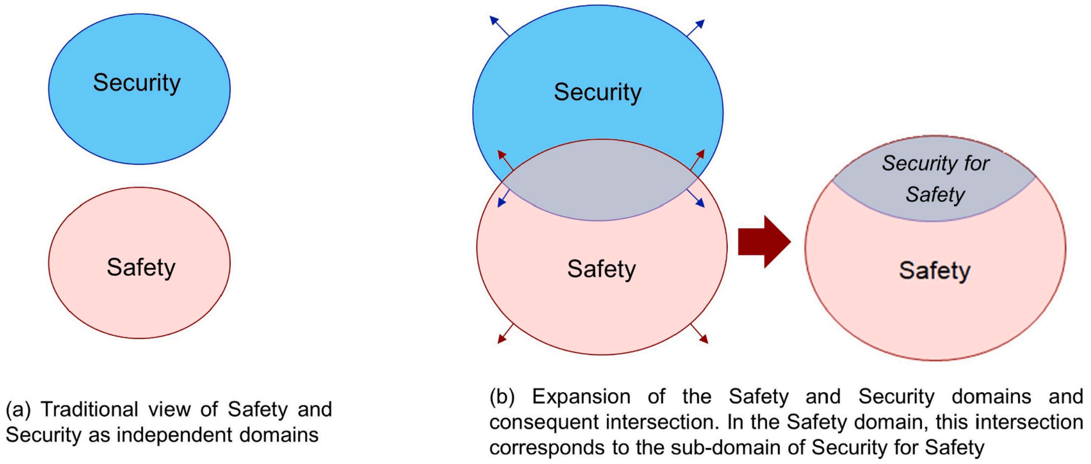
Referring to both malicious and clumsy users as causes of harm, Max Tegmark stresses how engineers should account for both of them as the sources of risk to be prevented by design in CPSs. In other words, responsible and intelligent design of CPSs should account for both safety and security issues. To make this type of analyses operational, some researchers and practitioners have recently proposed some methods that designers and engineers can use to perform risk analysis. An overview of these methods is provided below.

2.1. Existing methods for integrated safety and security analysis of CPSs

Standards that concern design, development, and operation/use of safety-related electrical/electronic/programmable electronic technologies, such as IEC 61508 (2010a,b) and their sector specific adaptions - e.g. ISO 26262 (2011), IEC 61511 (2016), and IEC 62061 (2010) -, only address security to a limited extent. Yet, some of them give reference to security standards like IEC 62443 (2010). IEC 62443 covers cyber security of industrial automation and control systems in general, with a few additional requirements added for what this standard regards as essential functions that cover safety systems. There are currently only a few initiatives that address the implications of security on safety and how safety requirements are formulated. An example of such initiatives is an early draft of IEC/TR 63069 (2019) being circulated with this topic as the main focus.

Although these standards provide relevant requirements and recommendations in their own domain, they do not provide fundamental theories and operational tools to achieve the combined goals they pursue (Lundteigen and Gran, 2019). New proposals in safety and security in other industries - such as ISO/TR 22100-4 (2018) in machinery and the standard under preparation ISO 21434 to complement ISO/PAS 21448 (2019) in autonomous vehicles -, could prove useful to bridge the gap of framing security for safety for CPSs.

Several researchers have performed surveys of safety and security



(a) Traditional view of Safety and Security as independent domains

(b) Expansion of the Safety and Security domains and consequent intersection. In the Safety domain, this intersection corresponds to the sub-domain of Security for Safety

Fig. 1. Incorporating security for safety in the domain of safety science.

analysis methods available in the literature (Chockalingam et al., 2013; Kriaa et al., 2015; Lyu, Ding, and Yang, 2019; Paul et al., 2016; Raspotnig and Opdahl, 2013). We have reviewed them and divided into two groups: one group is characterized by (1) clear and distinguishing novel theoretical foundations and (2) demonstrated applicability to real cases. The second group cannot be attributed these characteristics and the methods included in it belong rather to the academic realm, the practical use of which is unclear and which have not been subjected to any kind of validation. In Section 3 and Section 4, we mention the features of the particular theories and techniques falling into the second group.

In the remainder of this section, we focus on the first group of methods that we consider a representative sample against which we compare our method and tool.

The System-Theoretic Process Analysis (STPA) is a hazard identification method originally conceived for safety analysis (Leveson, 2004; Leveson and Thomas, 2018). STPA has now some derived alternatives known as STPA-Sec (Young and Leveson, 2013) and STPA-SafeSec (Friedberg et al., 2017) to include security into its original safety framework. However, STPA-Sec and its variation STPA-SafeSec only identify high-level socio-technical vulnerabilities independently from the analysis of how particular threats could target specific component vulnerabilities. Some researchers have criticized this perspective and argued that it could prove useful only for early design stages of the system lifecycle (Kriaa et al., 2015; Schmittner et al., 2016; Temple et al., 2017). In this view, we acknowledge there is a need to analyze also the particular vulnerabilities arising from the tactics employed by threat actors, including the vulnerabilities present in specific digital technologies and the interconnections that have proven to lead to new unsafe scenarios.

From a different perspective based on Unified Modelling Language (UML) diagrams, in his doctoral thesis Raspotnig (2014) proposed the combined harm analysis for safety and security of information systems (CHASSIS). CHASSIS provides insights into which vulnerabilities are more critical and have bigger influence in terms of their contribution to different complex scenarios. Still, this method only provides specification of functional and safety requirements looking at individual scenarios (Raspotnig et al., 2013). With respect to STPA-Sec, CHASSIS does provide insights on malicious actors and their tactics to exploit the system vulnerabilities. However, similarly to STPA-Sec, CHASSIS is also an approach that is suitable only for systems in early requirement and concept phase, where the level of resolution of the final system architecture is still unknown (Schmittner et al., 2015). In other words, these methods do not reach the level of detail needed to analyze the actual implementation of the system with the particular vulnerabilities associated to the components and technologies in place. Moreover, Schmittner et al. (2015) argue that both STPA-Sec and CHASSIS are highly dependent on expert knowledge, leading to different groups of analysts to obtain considerably different results. To assist the stakeholders during the risk analysis, STPA-Sec and CHASSIS do not provide a database of scenarios and recommendations of protection measures derived from lessons learned analysis, making the methods prone to completeness issues and diverging results.

In her doctoral thesis, Kriaa (2016) described the SCADA safety and security joint modelling method (S-cube). Building on previous work using a Boolean logic Driven Markov Processes (BDMP) formalism in industrial security scenarios (Kriaa et al., 2012), the S-cube method includes a taxonomy of attack vectors and a knowledge base of failure/attack steps and propagations (Kriaa et al., 2019). This knowledge base in S-cube addresses the dependency on expert knowledge mentioned for STPA-Sec and CHASSIS. Furthermore, S-cube offers a better scalability of the analysis that adapts to different levels of abstraction of the system (Kriaa et al., 2013). However, S-cube decouples the information flows in the digital control system from the analysis of energies controlled in the physical plant (Kriaa et al., 2019). This decoupling, aimed at keeping the analysis scope into the domain of the digital control system, may be a critical weakness to identify new and unknown risks that have important safety repercussions in the physical domain of the system.

Another effort in line with identifying cyber security vulnerabilities in the design of process plants is the sneak path security analysis (SPSA) method (Baybutt, 2003). SPSA is an extension of traditional sneak path analysis that aims at including cyber security. SPSA relies on topological diagrams of the system and a set of checklists to assist brainstorming sessions. These brainstorming sessions aim at identifying paths from threat sources to targets assets that may cause harm. While this method proves useful in simple system architectures in process plants, it lacks a generic system representation that would fit and facilitate analysis of complex CPSs.

In sum, we find two main limitations in these safety and security analysis methods. First, some methods tend to be suitable only for early system lifecycle phases or for simple system architectures. And second, some methods do not integrate the relationships between the information systems and the specific energies controlled by the CPSs in the physical world. To overcome these limitations, Carreras Guzman et al. (2020) introduced the concept coined Uncontrolled Flows of Information and Energy (UFoI-E). Acknowledging the benefits of this concept, we build upon it an integrated safety and security analysis method for CPSs that we refer to as the UFoI-E method. We dedicate the Section 3 to summarize the features and capabilities of this method. Afterwards, in Section 4 we describe in detail a new technique that we designed to make this method more systematic and to reinforce the performance of the analysis in terms of completeness.

3. Overview of the UFoI-E method

In this section, we provide a short overview of an integrated safety and security analysis method coined the Uncontrolled Flows of Information and Energy (UFoI-E). The two constituents of the method, the UFoI-E causality concept and system representation, have already been introduced in earlier publications; while the third constituent, the CyPHASS scenario builder, is presented for the first time below in this paper, Section 4. An encompassing representation of the UFoI-E method is given in Fig. 2.

In the following part of this section we briefly summarize the UFoI-E method.

3.1. UFoI-E causality concept

UFoI-E causality concept is a theoretical foundation of the UFoI-E method. Rooted in accident causation models (Hovden et al., 2010), this causality concept provides a terminology of safety and security and a model to abstract the causal chains in physical harm scenarios.

The core of the causality concept is very similar to the well-known Gibson's energy-barrier model (Rasmussen and Grønberg, 1997). However, the UFoI-E causality concept incorporates the domain of information into the model (Carreras Guzman and Kozine, 2018a, 2018a; Carreras Guzman et al., 2019a,b). This domain of information corresponds to the digital processes performed by information technologies and control systems whose goal is to control different functions related to information (e.g. obtain, store, compute, communicate). These information functions can potentially deviate from their intended performance and become uncontrolled flows of information (UFoI). These UFoI could propagate throughout the system and reach the domain of energy, becoming causes of uncontrolled flows of energy (UFoE). Furthermore, in the opposite direction, an UFoE could breach into the domain of information and lead to an UFoI, which in turn could be the precursor of another UFoE with physical harm implications. This bidirectional relationship between UFoI and UFoE constitutes the core of the UFoI-E causality concept. To prevent or mitigate the propagations between UFoI and UFoE, this causality concept introduces cross-domain influence barriers, which are design decisions or countermeasures to avoid propagation effects between the domains of the system.

In short, the UFoI-E causality concept emphasizes that cyber security threats and software flaws can also kill people and cause physical damages. In this view, the control of information flows becomes a critical requirement to prevent physical harm scenarios related to uncontrolled energies and materials.

3.2. System representation: The CPS master diagram

Risk analysts require not only documentation and descriptions of the system under analysis, but also a common model to represent these descriptions in a coherent way. In the words of P. L. Clemens, "we never analyze a system – we analyze only a conceptual model of a system" (Rausand, 2011). For example, typical HAZOP studies rely on system representations such as piping and instrumentation diagrams (P&IDs), FMEA studies usually rely on design drawings or functional block diagrams, and software safety analyses rely on diagrams such as the ones suggested by the Unified Modeling Language (UML).

In this context, the CPS master diagram is as a generic framework for the representation of CPSs to serve as a basis for safety and security analysis. The purpose of the CPS master diagram is to provide a common system conceptualization that multidisciplinary teams of experts can use to design diagrammatic representations of their particular systems under analysis. Therefore, experts can cooperate more efficiently while avoiding confusion in their communications. A detailed description is provided in Carreras Guzman et al. (2020).

The CPS master diagram has been applied to the representation of an autonomous surface vessel and to a driverless bulldozer prototype, proving its suitability to properly handle diverse types of CPSs for safety and security analysis (Carreras Guzman et al., 2019a; Carreras Guzman and Mezovari, 2019; Carreras Guzman et al., 2019b). Another example of application is given in Section 5.

In further work, researchers can compare the capabilities of other system representations in the context of risk identification. Some proposals include the STPA control structure (Leveson, 2011), UML diagrams as in CHASSIS (Raspotnig, 2014), the CORAS modelling language (Lund et al., 2011), among others. Furthermore, the combined use of the CPS master diagrams with these other system representations may prove useful to assist analysts in the visualization of different system vulnerabilities.

4. CyPHASS: A harm scenario builder

In this section, we introduce our main contribution to make the UFoI-E method an operational technique for safety and security analysis. The Cyber-Physical Harm Analysis for Safety and Security (CyPHASS) is a harm scenario builder designed to assist analysts working on CPS risk identification. We designed CyPHASS as a practical risk identification tool inspired by the UFoI-E causality concept and the CPS master diagram representation. Building on previous work of the UFoI-E method that used fault trees in a case study (Carreras Guzman et al., 2019a), CyPHASS is an extension of the bowtie method that is specifically tailored to the risk analysis of CPSs using an ontology of harm scenarios

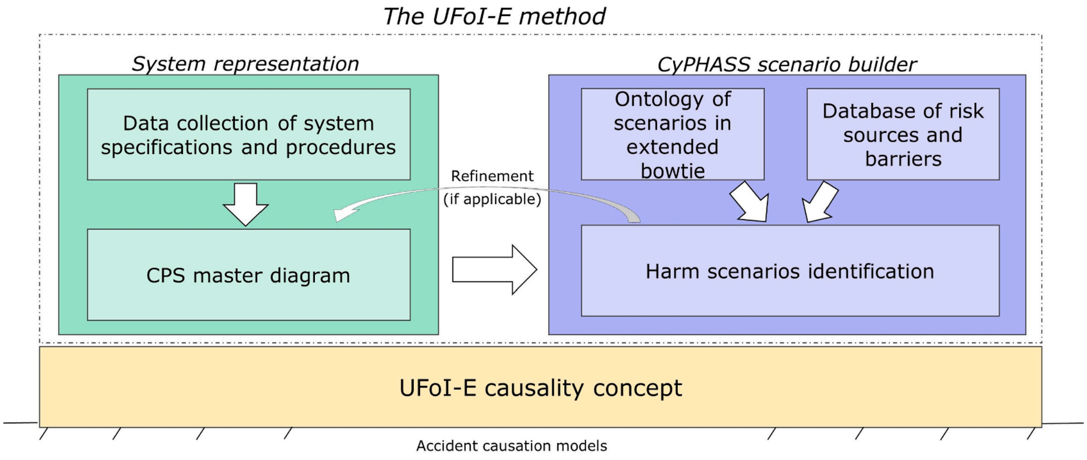

The UFoI-E method
System representation
Data collection of system specifications and procedures
CPS master diagram
Refinement (if applicable)
CyPHASS scenario builder
Ontology of scenarios in extended bowtie
Database of risk sources and barriers
Harm scenarios identification
UFoI-E causality concept
Accident causation models

Fig. 2. The UFoI-E method as an integrated approach for safety and security analysis.

and a populated database of checklists and guidewords.

In the following subsections, we describe the two main constituents on which the CyPHASS tool is built. First, we describe the ontology of scenarios in terms of an extended bowtie model. And second, we introduce the populated database of checklists embedded in CyPHASS. Serving as a practical toolkit, we share our software prototype of CyPHASS as an open-access supplemental material linked to this manuscript. Finally, we compare CyPHASS to other practical techniques for safety and security analysis.

4.1. *CyPHASS ontology of scenarios*

> "It is a golden rule of history that what looks inevitable in hindsight was far from obvious at the time"
> (Harari, 2015)

An ontology of scenarios provides a general framework to manage the complex task of identifying risk scenarios in CPSs. Its goal is to enable a systematic and comprehensive process of identification of risk scenarios in a sequence of stages. Risk scenario is a sequence of events starting from a threat/hazard and resulting in an event having the potential to harm.

We illustrate this ontology of scenarios as an extended bowtie model. Following the bowtie diagram convention, a top event is characterized by a set of causes to the left and a set of consequences to the right (de Ruijter and Guldenmund, 2016). As shown in Fig. 3, CyPHASS consists of a set of four sequential types of top events pictured as circles, each of them linked to a set of causes to its left and to a set of consequences to its right.

Reading the diagram backwards – i.e. from right to left – the first top event is the UFoE at the physical layer, which is the direct cause that may lead to ultimate safety consequences. According to the state of the art in the energy model, these UFoE events can also be releases of toxins and hazardous materials.

To the left of the UFoE events, we now find the sequential top events that correspond to system deviations at each layer of the CPS master diagram. Namely, the first top event (the UFoE) is preceded by the second top event corresponding to deviations at the physical layer (PL). These PL deviations are preceded by the third top event corresponding to deviations at the cyber-physical layer (CPL). And, in turn, these CPL deviations are preceded by the fourth top event corresponding to deviations at the cyber layer (CL). The CL and CPL deviations are possibilities of UFoI, whereas the PL deviations are possibilities of process variable deviations and functional deviations (PV-F). We expand on this difference in Section 4.2.

The four top events in CyPHASS are linked by event trees, each composed of three branches. These event trees characterize the propagation effects between the top events. To avoid these propagation effects, the event trees incorporate detection and response barriers. For each top event, if both the detection and response barriers are present and perform as planned, the scenario reaches a safe state. On the contrary, if either the detection or the response barrier is not present or gets breached, the scenario considers a propagation effect from a top event to the next.

Furthermore, the three top events to the left of the diagram are also linked to a set of direct causes. These direct causes are hazards and threats that menace each layer of the system. Namely, PL deviations could be directly triggered by physical hazards and threats targeting this layer. Cases of CPL UFoI and CL UFoI, in turn, could be triggered by either cyber hazards, cyber threats, physical hazards and/or physical threats targeting these layers directly. These hazards and threats can also originate from outside of the system, i.e. from the cyber environment (CE) or from the physical environment (PE).

Considering this possibility of direct hazards and threats, CyPHASS incorporates safety and security prevention barriers to eliminate or reduce the likelihood of these hazards and threats in the first place. For each identified threat or hazard, if the related prevention barrier is not present or gets breached, the scenario considers the occurrence of the related top event as a result. Again, if a top event occurs, it may subsequently propagate to the next top event as described in the previous paragraphs.

In practice, the bowties in CyPHASS allow risk analysts to conduct the identification of risk scenarios using a causal analysis methodology. Starting from the identification of UFoE as hazardous events, risk analysts can perform a step-by-step identification of sequential causes in a backward analysis. This systematic causal analysis supports the finding of propagation effects between the layers of the CPS, as well as common-cause sources that may lead to various UFoE scenarios. Subsequently, risk analysts can recommend safety and security barriers at different stages of the scenarios, providing layers of protection to prevent or mitigate propagation effects (DHS, NCCIC, ICS-CERT, 2016).

Therefore, for an organized stepwise process of the identification of risk scenarios, analysts can follow an algorithm as shown in Appendix A. This process of risk identification can be summarized in the following steps. In parenthesis, we point which parts of the CPS master diagram should be examined in each step.

Step 1: Identify the cases of UFoE that could lead to ultimate safety consequences (PL and PE)
Step 2: For each UFoE, identify the causes as PL PV-F deviations (PL)

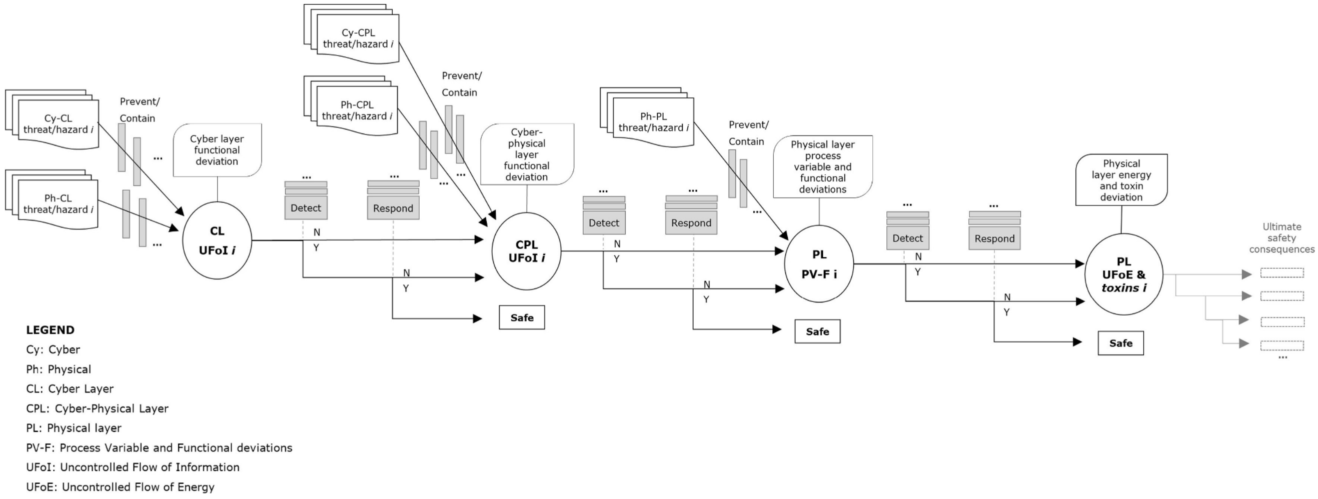

Cy-CPL threat/hazard *i*
Ph-CPL threat/hazard *i*
Cy-CL threat/hazard *i*
Ph-CL threat/hazard *i*
Prevent/ Contain
...
Cyber layer functional deviation
CL UFoI *i*
Detect
Respond
N
Y
Cyber-physical layer functional deviation
CPL UFoI *i*
Safe
Ph-PL threat/hazard *i*
Physical layer process variable and functional deviations
PL PV-F i
Physical layer energy and toxin deviation
PL UFoE & *toxins i*
Ultimate safety consequences
LEGEND
Cy: Cyber
Ph: Physical
CL: Cyber Layer
CPL: Cyber-Physical Layer
PL: Physical layer
PV-F: Process Variable and Functional deviations
UFoI: Uncontrolled Flow of Information
UFoE: Uncontrolled Flow of Energy

Fig. 3. CyPHASS ontology as extended bowties.

- **Step 2.1:** For each PL PV-F deviation, identify and recommend detection and response barriers (CL, CPL, PL)

Step 3: For each PL PV-F deviation, identify causes as physical hazards and threats H/T (PL, PE)

- **Step 3.1:** For each physical H/T, identify and recommend prevention barriers (PL)

Step 4: For each PL PV-F deviation, identify causes as CPL UFoI (CPL)

- **Step 4.1:** For each CPL UFoI, identify and recommend detection and response barriers (CL, CPL, PL)

Step 5: For each CPL UFoI, identify causes as cyber and physical H/T (CPL, CE, PE)

- **Step 5.1:** For each cyber and physical H/T, identify and recommend prevention barriers (CPL)

Step 6: For each CPL UFoI, identify causes as CL UFoI (CL)

- **Step 6.1:** For each CL UFoI, identify and recommend detection and response barriers (CL, CPL, PL)

Step 7: For each CL UFoI, identify causes as cyber and physical H/T (CL, CE, PE)

- **Step 7.1:** For each cyber and physical H/T, identify and recommend prevention barriers (CL)

Note that the detection and response barriers at different stages are not restricted to the layer in the CPS master diagram where the deviation occurs. Indeed, one of the benefits of using the CPS master diagram is the realization that the interacting layers of the system can provide various mitigation measures, even as redundant barriers. For example, a PL PV deviation such as “pressure too high” can be detected either by an operator at the PL hearing a sound alarm, by a safety-related sensor at the CPL reading the pressure value, or by a remote supervisor at the CL looking at an indicator colored in red in a HMI. In turn, each one of these entities can take a response action, arranged as a layers of protection strategy.

4.2. Database of checklists to build scenarios in CyPHASS

> “To understand the mechanisms of accidents and to develop accident prevention and control strategies, it is essential to know about and learn from past incidents”
> (Khan and Abbasi, 1999)

To build a database of scenarios, the CyPHASS tool incorporates generic checklists and guidewords. The approach we use to generate the checklists is systematic and built on the historical knowledge. That is to say, it is derived in a structured way from a lessons learned analysis of past events. It is systematic because it is strongly governed by the CyPHASS ontology of scenarios presented in Section 4.1. The ontology is then applied to break down and structure historic cyber-physical attacks in industrial facilities. The propagation paths of the attack scenarios become explicated so that measures preventing and/or mitigating them can be worked out. To generate a list of measures, which in fact constitute the checklists and guidewords, expert knowledge of subject matter specialists is invoked.

For illustrative purposes, in Table 1 we summarize a set of cyber-physical attacks against industrial plants used in our lessons learned analysis. This table is not the CyPHASS database of checklists, but a summary of some of the cases that we have used to populate the generic checklists. To identify the localization of each stage of the attack, we use the terminology of the ontology of scenarios in CyPHASS, as described in the previous subsection.

Table 1
Short summary of historic cyber-physical attacks in industrial facilities explained using the CyPHASS ontology.

Attack title	Country, Year	Scenario	Safety consequences	Reference
Maroochy Water Breach	Australia, 2000	(CE-CL) Access via unsecure radio communications → (CPL) reconfiguration of PLCs and deactivation of alarms → (PL) pumping stations uncontrolled	(PE) Release of one million liters of untreated sewage into a storm water drain from where it flowed to local waterways	(Slay and Miller, 2007)
Aurora Generator Test at Idaho National Laboratory	USA, 2007	(PE-CPL) Local access to programmable controllers and control network → (CPL) injection of malicious code → (PL) circuit breakers open and close out-of-sync	(PL) Damage and destruction of rotating equipment (generator, turbine, etc.)	(Zeller, 2011)
Baku-Tbilisi-Ceyhan pipeline attack	Turkey, 2008	(CE-CPL) Wireless access gained to security camera network + (PE-CPL) possibly physical access to field controllers → (CPL) penetration into control network → (CPL) command injections to controllers + (CL, CPL) suppression of alarms → (PL) pipe over pressurization	(PL) Pipeline rupture, explosion and fire	(Lee et al., 2014b)
Stuxnet	Iran, 2010	(CL) USB drive infected with Stuxnet inserted to Windows computer connected to network → (CPL) worm propagation to PLCs via local network → (PL) rotor speed manipulated + (CPL, CL) rotor speed masked to monitoring systems	(PL) Nuclear centrifuges ruined	(Langner, 2011)
German Steel Mill attack	Germany, 2014	(CE-CL) Spear phishing emails → (CL) access to corporate network → (CPL) penetration into plant network → (CPL, PL) multiple plant components damaged + (CL) HMI disrupted → (PL) blast furnace prevented to shut down	(PL) Serious damages to plant infrastructure	(Lee et al., 2014a)
BlackEnergy	Ukraine, 2015	(CE-CL) Spear phishing e-mails with BlackEnergy malware → (CL) access to business networks for extended period (more than 6 months) → (CL, CPL) use of VPNs to penetrate ICS network and connected HMIs and field devices → (PL) breakers opened uncontrolled + (CPL) malicious firmware injections + (CL) telephone DoS attack on the call center	(PE) Power outages that caused approximately 225,000 customers to lose power for several hours across various areas	(E-ISAC, 2016)
Crash Override (Industroyer)	Ukraine, 2016	(CE-CL) Presumably spear phishing e-mails → (CL, CPL) malware automatically maps out control systems and locates target equipment + (CL, CPL) malware records network logs that it can send back to hackers → (CPL, PL) launching payload modules reaching grid equipment	(PE) Power outage for a few hours in northern Kiev	(DRAGOS, 2017)
TRITON attack	Saudi Arabia, 2017	(CE-CL) Access to IT corporate network for extended period (almost 1 year) possibly via social engineering attack → (CPL) penetration into control network → (CPL) penetration reaching and controlling safety-instrumented systems	(PL) Attempt to deactivate safety-instrumented systems and cause a plant explosion	(Pinto et al., 2018)

CE: Cyber environment; CL: Cyber layer; CPL: Cyber-physical layer; PL: Physical layer; PE: Physical environment.

Other critical cyber-physical attack scenarios recently proven possible include cases against vehicles. In 2010, a team of researchers succeeded in remotely hacking cars while driving on the road in a controlled environment (Checkoway et al., 2011; Koscher et al., 2010). These cyber-physical attacks could start via the exploitation of vulnerabilities in a wide set of apparently non-safety related components such as CD players, Bluetooth connected components, and telematics systems using cellular radio networks **(CE threat → CL UFoI)**. Then, because the safety-critical networks and electronic control units (ECUs) are not completely isolated from the car's entertainment network, hackers could bypass network security in the car subnetworks and reach the safety-critical components **(CPL UFoI).** Researchers demonstrated how hackers could make the car ignore the driver's input, e.g. deactivating the brakes, activating the brakes in dangerous circumstances, or turning off the engine **(PL deviation)**. Subsequently, the attackers could control the speed of the vehicle at will **(PL UFoE)** and compromise the safety of the passengers and the surroundings of the vehicle **(PL-PE safety consequence)**. Moreover, multi-stage attacks could combine different attack vectors, making the dashboard display wrong information to confuse the drivers (e.g. wrong speed in speedometer) and even erasing any evidence of the cyber intrusion afterwards **(CL UFoI)**.

Similar vulnerabilities that could have led to cyber-physical attacks against Jeep vehicles were unveiled by Miller and Valasek at the Black Hat USA 2015 conference (Miller and Valasek, 2015). In view of these facts, the National Highway Traffic Safety Administration (NHTSA) of the United States released a guidance for best practices in cyber security of modern vehicles (National Highway Traffic Safety Administration, 2016). Beyond cars, cases in other transportation systems also show how of cyber-physical attacks could hijack the control of ships (DNV GL, 2016; Torkildson, 2018) and unmanned aerial vehicles (UAVs) (Plioutsias et al., 2018; Yampolskiy et al., 2012).

Considering these lessons learned, and incorporating expert knowledge from safety and security analysts, we have developed a CyPHASS prototype with generic checklists. As shown in Fig. 4, these checklists are generic functional deviations and guidewords related to each element of the CyPHASS bowtie. Using widespread software tools such as worksheets, risk analysts can navigate throughout the CyPHASS bowtie and systematically assess the CPS under analysis using these checklists. For each case in the set of checklists and guidewords, the analysts should ask the question: "Is this possible in my system?" As previously mentioned in Section 4.1, the analysis can be conducted in a stepwise process as described in the algorithm in Appendix A.

Serving as a practical toolkit, we share the CyPHASS databases of checklists as an open-access supplemental material linked to this manuscript. The checklists include examples of top events, i.e. CL UFoI, CPL UFoI, PL deviations and PL UFoE. To avoid the propagation effects between top events, we provide checklists of detection and response barriers. Furthermore, as causes of each type of top event, we include checklists built from examples of hazards and threats. And finally, for different types of threats and hazards, we include checklists of prevention barriers to eliminate them or reduce their likelihood.

Expanding on the PL deviations, we stress that CyPHASS is not a typical cyber security tool aimed at ensuring the generic CIA security properties. Because the scope is safety, the concept of security for safety encompasses those security properties with the direct potential for safety consequences related to the physical world. In this regard, for the top event PL deviations we provided a set of two checklists. The first checklist is called process variable (PV) deviations and the second checklist is called functional (F) deviations. This separation into two set of checklists was an interesting conclusion following a series of workshops with safety experts discussing the emerging challenges to assess the PL of CPSs. Below, we explain the reasoning behind this conclusion.

The PV checklist builds on the HAZOP technique to provide a list of process variables and a set of deviation guidewords (Dunjó et al., 2010). We recommend this checklist for the context of CPSs controlling energy sources and hazardous materials using mechanical and electrical equipment, where HAZOP has proven a very successful approach (Taylor, 2017). These cases include the context of process plants, power plants, oil and gas facilities, water and wastewater plants, nuclear plants, among others. Typical examples include PV such as "pressure" and "speed" and HAZOP deviations such as "too high" and "too fast".

As an alternative to the PV checklist, the F checklist provides guidance to perform a functional assessment of the PL of the system. We provide examples of functions, functional parameters and functional deviations that may lead to UFoE. This approach is a version of the functional hazard analysis that has proven a powerful technique to analyze software-intensive systems (Ericson, 2005). We recommend these checklists for the contexts of CPSs that perform a wide set of functions to control energy in the physical world, particularly if these systems are novel or in a prototype phase. This context includes autonomous vehicles, drones, robotics and machinery, smart medical devices, military defense systems, among others. This context also includes particular safety systems providing safety functions on demand. Typical examples include functions such as "launch" or "stop", functional parameters such as "correctness" or "timing", and their respective functional deviations such as "omission" or "delay".

Finally, both the PV and the F checklists can be used in combination, providing alternative lenses to the analysts to assess the same system and avoid missing PL deviations. By providing tailored checklists for the PL deviations and PL UFoE, CyPHASS overcomes the limitations of traditionally independent cyber security analysis, safety analysis of control systems and safety analysis of physical systems. Overcoming these limitations is not only a fine integration, but a requirement to conduct risk identification in the cyber-physical domain. In CPSs, the cyber and the physical layers are tightly interconnected, and independent analyses may no longer be a sufficient strategy.

4.3. CyPHASS bowtie compared to similar techniques in the literature

When compared to traditional risk identification techniques in safety analysis, there is a clear connection between the extended bowtie model in CyPHASS and the traditional techniques of fault tree analysis (FTA) and event tree analysis (ETA). Indeed, the risk sources and the prevention barriers on the left side of the top events can be represented using the conventions of fault trees, whereas the detection and response barriers on the right side of the top events can be represented using the conventions of event trees (de Ruijter and Guldenmund, 2016). Alternatively, both sides of the bowtie can be represented in the form of safety-barrier diagrams, where the logic gates of the fault trees and event trees are simplified using sequential barriers (Duijm, 2009).

When compared to risk identification techniques in security analysis, there is a direct connection between CyPHASS and attack trees (Schneier, 1999). In line with the method proposed by Nai Fovino et al. (2009), the direct H/T risk sources in CyPHASS can be conceived as the integration of fault trees and attack trees.

For the domain of industrial control systems (ICSs), Abdo et al. (2018) proposed an integration of attack trees into bowtie analysis that is similar to our approach in CyPHASS. The main difference is that CyPHASS provides an extended bowtie ontology of harm scenarios that consists of four consecutive top events. Based on the CPS master diagram, this ontology of harm scenarios - combined with the populated database of checklists and guidewords - provides a comprehensive framework to conduct a stepwise analysis able to trace propagation effects among the layers of the CPS. These features in CyPHASS provide stronger capabilities in the qualitative identification of safety and security scenarios. However, regarding quantitative features, the model in Abdo et al. (2018) provides more guidance for estimation of likelihood and consequences of safety and security events. For further work, this comparison provides good motivation to explore a potential integration of the capabilities of these two methods.

Finally, we can derive an analytical discussion between CyPHASS and the Cyber Kill Chain framework of the Lockheed-Martin corporation

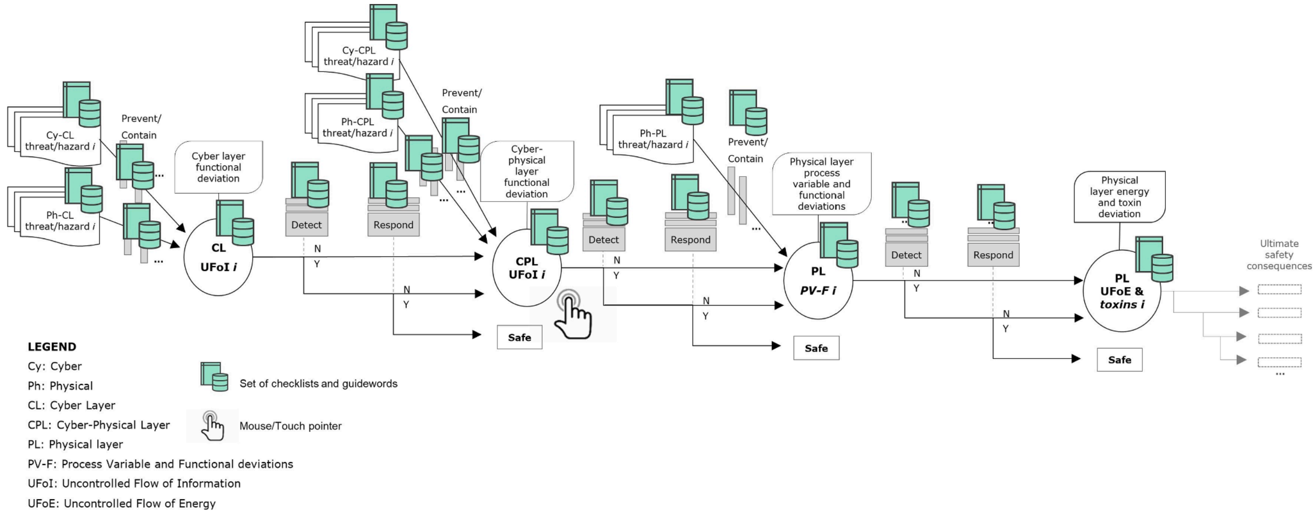

Cy-CL threat/hazard *i*
Ph-CL threat/hazard *i*
Prevent/ Contain
Cyber layer functional deviation
CL UFoI *i*
Detect
Respond
N
Y
Cy-CPL threat/hazard *i*
Ph-CPL threat/hazard *i*
Prevent/ Contain
Cyber-physical layer functional deviation
CPL UFoI *i*
Safe
Ph-PL threat/hazard *i*
Prevent/ Contain
Physical layer process variable and functional deviations
PL *PV-F i*
Physical layer energy and toxin deviation
PL UFoE & *toxins i*
Ultimate safety consequences
LEGEND
Cy: Cyber
Ph: Physical
CL: Cyber Layer
CPL: Cyber-Physical Layer
PL: Physical layer
PV-F: Process Variable and Functional deviations
UFoI: Uncontrolled Flow of Information
UFoE: Uncontrolled Flow of Energy
Set of checklists and guidewords
Mouse/Touch pointer

Fig. 4. Illustration of CyPHASS tool with databases of generic checklists and guidewords.

(Hutchins et al., 2011). The Cyber Kill Chain subdivides cyber security attacks into seven stages: reconnaissance, weaponization, delivery, exploitation, installation, command and control, and actions on objective. Because the first two stages of the Cyber Kill Chain occur outside the system, these stages are beyond the scope of CyPHASS. In the terminology of CyPHASS, the delivery and exploitation stages correspond to the description of the threat, the installation and command and control stages correspond to the description of the UFoI, and the actions on objective correspond to the consequences downstream in the bowtie.

Arguably, the Cyber Kill Chain framework provides a more detailed description of cyber security attacks in terms of intrusion detection and defense of computer systems. Nevertheless, we highlight that CyPHASS considers both cyber and physical security attacks in an integrated analysis and have the specific scope to identify cyber-physical attacks against CPSs that could lead to UFoE events. In other words, the Cyber Kill Chain does not explicitly describe the physical implications of the cyber-attack and its relationship with safety, which CyPHASS considers as the mechanisms linking PL PV-F deviations to PL UFoE and their consequences. Therefore, we argue that the CyPHASS bowtie ontology of scenarios provides sufficient structure to illustrate the relevant stages of cyber-physical attacks for risk identification in CPSs and to propose suitable prevention, detection and response barriers, including intrusion detection systems.

5. Case study: Safety and security analysis of an industrial control system

In this section, we demonstrate an application of the UFoI-E method to analyze an industrial control system in the nuclear sector. We used the CPS master diagram for system representation and, based on this representation, we identified a set of harm scenarios using the CyPHASS ontology and its databases of checklists. Finally, we identified and recommended prevention, detection, and response barriers as intervention measures for each scenario identified in CyPHASS.

The system under analysis was the Halden Safety Fan Enclave, an on-demand emergency ventilation system installed in the Test Enclave laboratory of the Institute for Energy Technology (IFE) in Norway (Simensen et al., 2019). In nuclear reactors, safety fans maintain negative pressure in the reactor building in order to prevent radiation releases.

5.1. System representation with a tailored CPS master diagram

In a one day workshop session, two risk analysts and one facilitator mapped a CPS master diagram of the system specifications of the Halden Safety Fan Enclave. As background, all participants were previously familiarized with the documents and diagrams described in the system specifications. The facilitator ensured that all participants had these system specifications in printed format during the workshop.

Using the generic CPS master diagram as a template, the workshop participants subdivided the components of the Halden Safety Fan Enclave according to their position in the cyber, cyber-physical, and physical layer. All the workshop participants worked collaboratively in a wide screen using a basic diagram application. Furthermore, the participants connected the multi-layered interactions between the components according to the respective information flows and energy flows. Fig. 5 illustrates this system representation of the Halden Safety Fan Enclave. We refer to this system representation as the tailored CPS master diagram of our particular system.

According to the system architecture, an air pressure differential sensor measures the inbound and outbound pressures of the centrifugal fan. These sensor measurements are transmitted to a programmable logic controller (PLC) that monitors the pressure differential, calculates the change in pressure and establishes the needed change in the rotation speed of the fan. Subsequently, the PLC issues a command to an actuator (power inverter) to adjust the power output frequency controlling the fan. A power supply unit (PSU) is the source of 220 V powering the PLC and the actuator. Overall, this control loop is monitored by a supervisory control station; a computer is connected to the PLC to report the status of the controller and display the information to a human supervisor using a HMI.

In this tailored CPS master diagram we do not disclose the technical details of the components installed in the Halden Safety Fan Enclave. However, we highlight the use of the internet protocol suite TCP/IP in the local area network (LAN) connecting the PLC to the supervisory admin computer at the cyber layer. Moreover, at the cyber layer this LAN is connected with a higher level enterprise network. These indications in the tailored CPS master diagram served to identify some cyber security threat scenarios in CyPHASS.

Also in the tailored CPS master diagram, the workshop participants identified the human agents interacting with the system from cyber and physical interfaces. At the cyber layer, the remote supervisor is the most evident case because it was explicitly mentioned in the specifications of the system architecture. This remote supervisor monitors the status reports from the PLC and is able to intervene via the computer connected to the PLC, either editing the code in the control algorithm or remotely calibrating the pressure sensors connected to the PLC. Using the generic CPS master diagram, the workshop participants were able to identify other human agents interacting with the system in the tailored CPS master diagram.

- At the cyber layer, enterprise managers are also connected to the LAN.
- At the physical layer, field operators have physical access to the inverter in maintenance and repair operations.
- At the cyber environment, maintenance staff and vendor staff may access the admin PC at the cyber layer or the PLC at the cyber-physical layer for software update purposes. Furthermore, remote vendor staff could connect to the enterprise TCP/IP network via the Internet.
- At the physical environment, some person may try to gain physical unauthorized access to the Enclave room where the Halden Safety Fan is located.

All the previously identified human agents may be benevolent actors and they could make unintentional mistakes, violations or slips that alter the behavior of the system (Reason, 1990). However, these actors can also have malicious intentions, acting as remote hackers (cyber environment), external saboteurs (physical environment), or malicious insiders (cyber and physical layers).

Finally, the workshop facilitator printed the tailored CPS master diagram in large format, so all workshop participants could visualize it and refer to it during the following risk identification session.

5.2. Identification of harm scenarios using CyPHASS

Using the tailored CPS master diagram as system representation, the one day workshop continued with a risk identification session using CyPHASS. As mentioned in Section 4.1 and Appendix A, CyPHASS offers a systematic process that can be performed in a series of steps using the ontology of scenarios in bowtie structure and a database of generic checklists and guidewords.

Using a wide screen and a spreadsheet application, all the workshop participants used the generic checklists to collaboratively identify the UFoE, the PV deviation, the PL hazards/threats, and all the barriers associated to the physical layer. Due to time limitations, all the participants addressed only the physical layer of the system collaboratively. Afterwards, the facilitator continued with the rest of the risk identification in a second day of work, addressing the cyber-physical and the cyber layers of the system. In this sense, the tailored CPS master diagram and the CyPHASS ontology and database provided sufficient guidance to continue the analysis in a consistent way. Finally, in a subsequent meeting the facilitator illustrated the final results to the rest of the workshop participants.

Fig. 6 summarizes the number of scenarios identified using CyPHASS. We counted a total of 12 independent scenarios of UFoI-E, 25 total cases of risk sources (hazards/threats), and 49 total barriers to protect the system as prevention, detection and response actions across the layers of the system. Some barriers were already present (P) in the system, while other barriers were recommended (R) as a result of the risk identification session.

To illustrate these results, the following paragraphs describe in detail three independent scenarios of UFoI-E identified with their associated hazards/threats and barriers. Each one of these independent scenarios has its risk sources at a different stage of the bowtie, serving as representative examples of each type of scenario in the CyPHASS ontology. Conveniently for illustration purposes, all three independent scenarios are associated to a unique **UFoE**, which is a **release of radioactivity**.

5.2.1. First illustrative scenario (PL PV deviation caused by PL risk sources)

PL PV deviation: This scenario leads to the radiation release (UFoE) due to a deviation in the physical layer of the system. Using the HAZOP guidewords of PV deviations, the radiation release can be caused by a PV deviation in the safety fan. Namely, the PV is **fan rotation speed** and the deviation is **too low or opposite direction**. As shown in the tailored CPS master diagram, this PV deviation would lead to low air flow from the fan and consequently to loss of assurance of negative pressure in the reactor building.

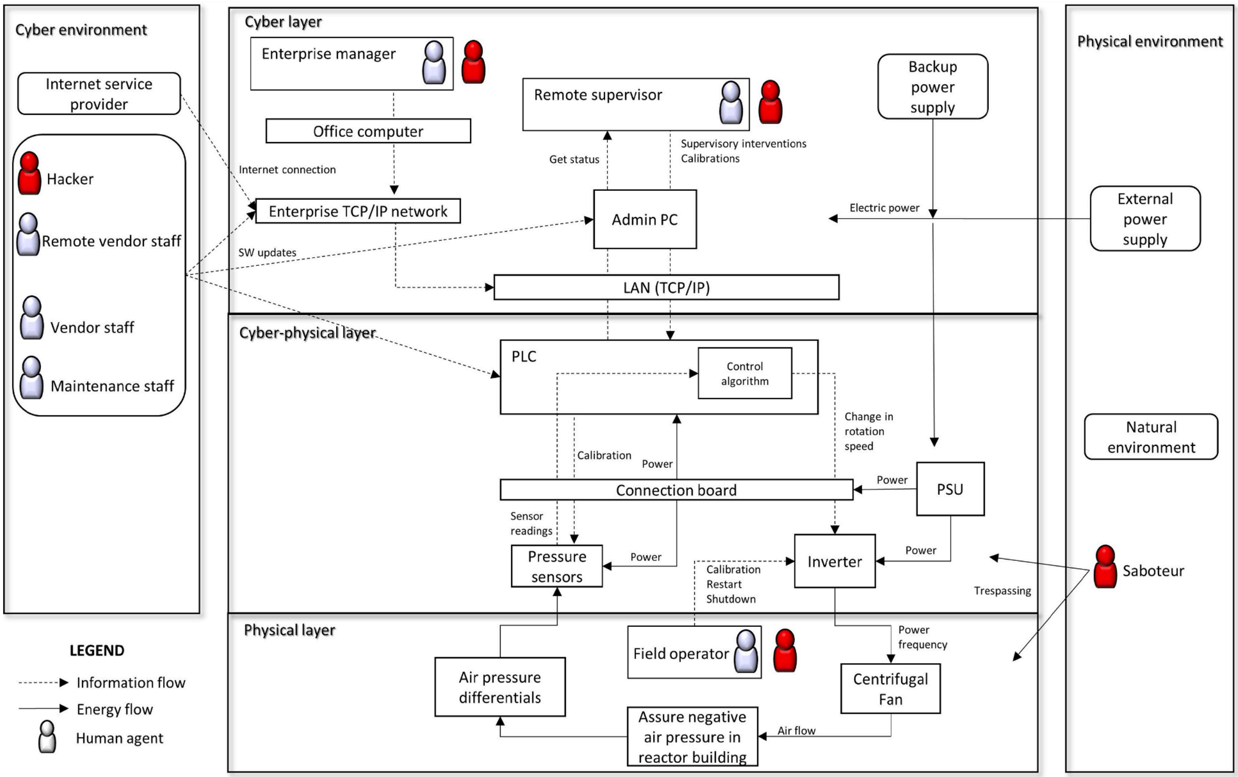


Fig. 5. Tailored CPS master diagram of the Halden Safety Fan Enclave.

PL detection and response barriers: Using the CyPHASS database to redirect the scenario into a safe state without reaching the UFoE, one of the detection barriers recommended was a visual detection and sound alarm at the admin PC alerting the remote supervisor at the cyber layer. This detection barrier can make use of the pressure sensor readings (present in the system) and an additional sensor measuring the fan rotation speed (recommended to add). As a response barrier, the supervisor can activate an emergency plan (recommended to add) and initiate a damage repair process that, if executed on time, would avoid the UFoE.

PL threat/hazards: To the left of the PV deviation in the bowtie, direct hazards/threats to the PL are causes of the PV deviation. Component failures and power loss are typical hazards identified that could be further analyzed in-depth. Scenarios of sabotage were also present as typical security threats to the PL. Furthermore, the CyPHASS database provided a case of human error from a field operator, i.e. physical components of the fan misplaced or wrongly connected after a maintenance/repair operation.

PL prevention barriers: For each PL hazard/threat, CyPHASS recommends the allocation of prevention barriers. Here we mention some of the ones identified. The component failures can be prevented with periodical maintenance plans (present in the system). The sabotage scenarios can be prevented with physical security measures such as access restriction with security cards, security cameras, and burglar alarms, among others (present in the system). And the maintenance/repair errors can be prevented not only with inspection after the maintenance/repair work, but also with additional sensors of fan integrity, alignment and calibration that the PLC can check automatically before restarting operations (recommended to add).

5.2.2. *Second illustrative scenario (CPL UFoI caused by CPL risk sources)*

CPL UFoI: This scenario continues tracing backwards the causal chain from the previously identified PL PV deviation, going further to the left in the CyPHASS bowtie. At this point, one of the CPL UFoI identified was **control logic in PLC reprogrammed maliciously**. This is a typical case of logic integrity violation in the CyPHASS database that integrates the lessons learned from past events such as the Stuxnet attack.

CPL UFoI detection and response barriers: This CPL UFoI can be detected indirectly by an additional software module in the PLC, or by parallel system independent from the PLC. Whenever possible, the independent solution is recommended, because an informed attacker may also reprogram the additional software module in the compromised PLC. The detection barrier could detect the reprogramming of the control logic indirectly by using combined feedback coming from the pressure sensors, fan speed sensors, and sensors of fan physical integrity. From this combined feedback, the parallel system could determine that the control algorithm in the PLC is issuing a command that is out-of-range from the normal operations and could subsequently send an alarm to the remote supervisor at the cyber layer. As response barriers, the remote supervisor - or the parallel system itself - could (1) activate a redundant safety fan system, (2) shutdown the compromised system, (3) isolate the compromised network for infection containment, and (4) initiate an infection removal process. If executed on time, these responses would avoid the PL PV deviation.

CPL hazards/threats: As direct causes of the CPL UFoI, one direct hazard/threat to the CPL identified was the case of a benevolent (or malevolent) agent connecting an infected device to the PLC. This case is possible considering that the PLC has a Modbus TCP interface. From the tailored CPS master diagram, a malicious insider could infect the PLC intentionally, but also vendor staff or maintenance staff could potentially execute this threat unintentionally.

CPL prevention barriers: Using the CyPHASS database, the allocation of authentication and privilege protocols can prevent unauthorized access to the PLC from a malicious insider. For the other cases of unintentional infection, an in-depth penetration testing analysis is needed to study how to prevent unsecure connections to the PLC by design.

5.2.3. *Third illustrative scenario (CL UFoI caused by CL risk sources)*

CL UFoI: This scenario continues tracing backwards the causal chain from the previously identified CPL UFoI. At this point, we identified an AND-gate combination of two CL UFoI. Namely, the **admin computer and its HMI are corrupted with malware** AND **the malware propagates throughout the LAN with the aim to gain access to the PLC**. This scenario is the longest chain in the CyPHASS bowtie that includes the sequential stages of CL UFoI → CPL UFoI → PL PV → PL UFoE. Note that CyPHASS provides a layers of protection strategy to mitigate each stage of propagation with detection and response barriers.

CL UFoI detection and response barriers: These CL UFoI can be detected by an intrusion detection system (IDS). This IDS could alert the remote supervisor via HMI notifications showing that the admin computer is infected. Furthermore, the recommended use of honeypots can redirect malicious traffic in the LAN, detecting and blocking the malware propagation. As response barriers, the analysis recommended (1) protocols for emergency communication with plant management, (2) network isolation for infection containment, (3) activation of redundant computational and network resources to supervise the fan system, and (4) antivirus diagnosis and repair to remove the malware.

CL hazards/threats: As causes of the CL UFoI, a wide set of hazards/

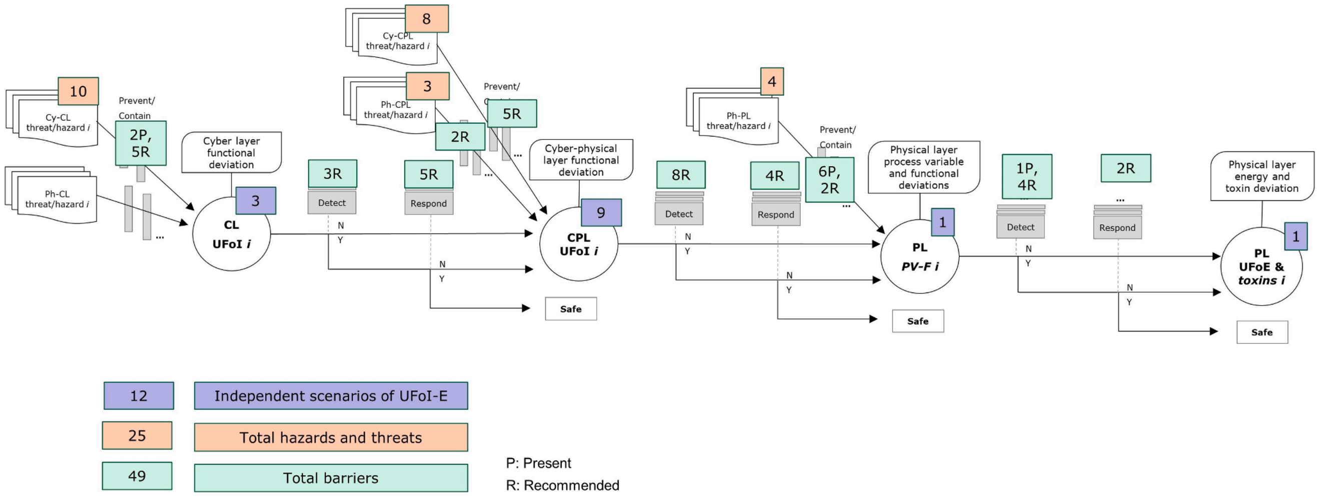


Fig. 6. Summarized results of harm scenarios identified using CyPHASS.

threats were identified using the CyPHASS database. These risk sources include the direct infection of the admin computer via the connection of an infected device (e.g. USB flash drive). The remote supervisor, the vendor staff, or the maintenance staff can do this unintentionally, or a malicious insider can do this intentionally. However, these are not the only possibilities. Because the CPS master diagram shows that the LAN is connected to an enterprise network and the networks communicate via TCP/IP, a hazard/threat may originate from the enterprise network connected to the Internet. These risk sources include cases of social engineering such as spear-phishing via email or man-in-the-middle attacks actively eavesdropping communications between enterprise managers and the remote supervisor. In each of these cases, an injected malware or eavesdropping of passwords can lead to the CL UFoI.

CL prevention barriers: To prevent this wide set of CL hazards/threats, several prevention barriers were recommended. An in-depth penetration testing analysis can prevent unsecure device connections to the admin computer by design. Although the LAN and the enterprise network is already separated by a firewall, an additional IDS could monitor the traffic and prevent a wider range of unauthorized access and misuse cases. Furthermore, the use of updated antivirus software should be enforced in the admin PC and in the enterprise computers. Finally, training of personnel against social engineering attacks and robust spam filters can reduce the likelihood of spear-phishing attacks.

5.3. Discussion

Our case study demonstrated that the UFoI-E method can be applied successfully in workshop sessions to conduct risk identification. First, the CPS master diagram provided a practical template for collaborative system representation. And second, CyPHASS ensured a systematic process of risk identification using the bowtie stages and their respective checklists. Furthermore, provided that all the workshop participants agree on a tailored CPS master diagram, the UFoI-E method also offered sufficient guidance to the workshop facilitator to complete the risk identification process with CyPHASS after the workshop session concluded. Particularly, we argue that the checklists in CyPHASS mitigate the issue of diverging results among risk analyst groups.

However, checklists being supporting tools for threat and hazard scenarios identification can be incomplete and biased, which reduces the predictive power of risk analyses. A way to overcome this possibility is to let users to extend the checklists as new knowledge becomes available.

From the workshop session, we also collected feedback from the participants to improve the CyPHASS database. Especially, the participants provided the idea to include two parallel checklists to analyze the physical layer of the system. As a result, we enhanced the CyPHASS database with the checklists for PV-F deviations, as described in Section 4.

In terms of the scenarios illustrated in Section 5.2, we selected the second and third illustrative scenarios to describe in detail some cases of security for safety in the analysis using CyPHASS. However, we highlight that other scenarios also described purely unintentional risk sources that reach all the way back to CL UFoI in the causal chain.

To summarize an example, one scenario is the possibility of HMI indicators ambiguous or hidden (CL hazard) → Remote supervisor error in parameter modification (CL UFoI) → Setpoints and sensor calibration corrupted (CL UFoI) → Fan rotation speed too low or opposite direction (PL PV) → Release of radioactivity (PL UFoE). To prevent, detect and respond to these scenarios, additional safety barriers at the CL and the CPL were recommended.

To further improve the method and enhance its predictive power, we have conducted two independent analyses of an autonomous surface vessel: one analysis applied the UFoI-E method while the other applied the STPA-SafeSec. The results of the study are published in (Carreras Guzman et al., 2021). This study exhibited strong and weak points of each method, their labor intensity and usability. The both methods learned from each other, the used vocabulary has been unified and checklists updated.

6. Conclusions

This paper demonstrated an integrated method to conduct safety and security analysis of CPSs. This is our response to the call to achieve a comprehensive analysis of the interactions among the CPS layers and system surrounding environments, preventing physical harm from a combined safety and security risk analysis as no current method had sufficiently achieved it so far (Zio 2018).

By having introduced the CyPHASS harm scenario builder, we can state now that the theoretical part of the UFoI-E method has been complemented by the missing operational tool. The practical relevance and usability of the method have been demonstrated by a number of cases: one demonstrated here, while other described in Carreras Guzman et al. (2019a) and Carreras Guzman et al. (2021).

We summarize the contributions of this paper in four main groups.

First, this paper explained the need to integrate cyber-physical security in the domain of safety science. We emphasized that, in CPSs, cyber security threats and software flaws can also kill people and cause physical damages. Therefore, this paper described an extended domain of safety science to incorporate the emerging subdomain of “security for safety”.

Second, we provided an overview of the UFoI-E method. The CyPHASS harm scenario builder is the third building block of the whole method that along with the two earlier developed blocks - the UFoI-E causality concept and the CPS master diagram - makes the method complete and a practical tool for a systematic safety analysis of CPSs that includes security.

Third, we explained in detail how CyPHASS is our main contribution to make the UFoI-E method a systematic technique that researchers and practitioners can use for risk identification. We presented the CyPHASS ontology of scenarios as an extended bowtie model, and described the populated databases of checklists and guidewords available in a prototype version. CyPHASS aims at bridging the gap between traditional safety analysis of control systems and traditional safety analysis of physical systems. Moreover, it aims at overcoming the limitations of traditional cyber security analyses conducted independently from the physical processes controlled by the software. We argued that overcoming these limitations is not only a fine integration, but a requirement to conduct risk identification in the cyber-physical domain.

And forth, we demonstrated an application of the UFoI-E method (and CyPHASS as a part of it) in an illustrative case study in the industrial sector.

We are certainly conscious that no one integrated security and safety analysis method can exhaustively identify harm scenarios in CPSs. Neither can the method presented in this paper. Nevertheless, it provides a different approach that has proven workable, that enriches the existing set of methods, provides a novel theoretical and practical contribution, can further learn from new knowledge and other approaches and be improved.

A common weakness of the whole family of integrated security and safety analysis methods is that they all employ a reductionist approach to the identification of harm scenarios. That is to say, the scenarios are predicted by examination of system’s individual parts. This way, emergent hazard scenarios cannot be identified, as they emerge from a system without arising from any part of the system alone, but because of interactions between parts. They are often manifested in complex, highly coupled, systems, possibly in catastrophic ways. This, in our view, is a remaining challenge of a great importance that is left for future research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We express our gratitude to Dr. Robert Taylor - Emeritus Research Fellow from Technical University of Denmark - for his valuable contributions to the lessons learned analysis and for his participation in the case study. We also thank Bjørn Axel Gran and André A. Hauge from the Institute for Energy Technology (IFE) for providing the system specifications of the Halden Safety Fan Enclave.

Appendix A. Algorithm to build scenarios using CyPHASS

List of acronyms:
CyPHASS: Cyber-Physical Harm Analysis for Safety and Security
UFoE: Uncontrolled Flows of Energy
PV-F: Process Variable - Functional
UFoI: Uncontrolled Flows of Information
PB: Prevention Barrier
DB: Detection Barrier
RB: Response Barrier
H/T: Hazards and Threats
PL: Physical Layer
CPL: Cyber-Physical Layer
CL: Cyber Layer

```
START
1. Identify PL UFoE
2. For PL UFoE i = 1 to n do
    3. Identify causes as PL PV-F deviations and potential combinations
[end of for 2.]

4. For PL PV-F deviations i = 1 to n do
    5. Identify PL PV-F deviations DBs
    6. Identify PL PV-F deviations RBs
    7. Identify causes as PL H/Ts and potential combinations
[end of for 4.]

8. For PL H/Ts i = 1 to n do
    9. Identify PL PBs
[end of for 8.]

10. For PL PV-F deviations i = 1 to n do
    11. Identify causes as CPL UFoI and potential combinations
    12. Identify potential combinations of PL H/Ts AND CPL UFoI
[end of for 10.]

13. If new PL H/Ts identified in 12, go to 8.

14. For CPL UFoI i = 1 to n do
    15. Identify CPL UFoI DBs
    16. Identify CPL UFoI RBs
    17. Identify causes as CPL H/Ts and potential combinations
[end of for 14.]

18. For CPL H/Ts i = 1 to n do
    19. Identify CPL PBs
[end of for 18.]

20. For CPL UFoI i = 1 to n do
    21. Identify causes as CL UFoI and potential combinations
    22. Identify potential combinations of CPL H/Ts AND CL UFoI
[end of for 20.]

23. If new CPL H/Ts identified in 22, go to 18.

24. For CL UFoI i = 1 to n do
    25. Identify CL UFoI DBs
    26. Identify CL UFoI RBs
    27. Identify causes as CL H/Ts and potential combinations
[end of for 24.]

28. For CL H/Ts i = 1 to n do
    29. Identify CL PBs
[end of for 28.]
STOP
```

Appendix B. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ssci.2021.105458.

References

Abdo, H., Kaouk, M., Flaus, J.M., Masse, F., 2018. A safety/security risk analysis approach of Industrial Control Systems: A cyber bowtie – combining new version of attack tree with bowtie analysis. Comput. Sec. 72, 175–195. https://doi.org/10.1016/j.cose.2017.09.004.

Amundrud, Ø., Aven, T., Flage, R., 2017. How the definition of security risk can be made compatible with safety definitions. Proc. Instit. Mech. Eng., Part O: J. Risk Reliab. 231(3), 286–294. Doi: 10.1177/1748006X17699145.

Aven, T., 2014. What is safety science? Saf. Sci. 67 (0925), 15–20. https://doi.org/10.1016/j.ssci.2013.07.026.

Baybutt, P., 2003. Sneak Path Security Analysis (SPSA) for Industrial Cyber Security. Primatech Inc.

Carreras Guzman, N.H., Kozine, I., 2018a. Identifying flows of information and energy in cyber-physical systems: A framework for safety risk analysis. In: 4th Society for Risk Analysis Nordic Conference. Stavanger.

Carreras Guzman, N.H., Kozine, I., 2018b. Uncontrolled flows of information and energy in cyber-physical systems. Retrieved November 19, 2018, from http://www.esrahomepage.eu/filehandler.ashx?file=16438.

Carreras Guzman, N.H., Kufoalor, D.K.M., Kozine, I., Lundteigen, M.A., 2019. Combined safety and security risk analysis using the UFoI-E method : A case study of an autonomous surface vessel. In: Beer, M., Zio, E. (Eds.), 29th European Safety and Reliability Conference (ESREL 2019). Hannover, pp. 4099–4106.

Carreras Guzman, N.H., Mezovari, A.G., 2019. Design of IoT-based cyber – physical systems: A driverless bulldozer prototype. Information 10 (11). https://doi.org/10.3390/info10110343.

Carreras Guzman, N.H., Mezovari, A.G., Yan, Y., Petersen, M.L., 2019. An IoT-Based Prototype of a Driverless Bulldozer. In: 2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS), pp. 291–296. IEEE. Doi: 10.1109/DCOSS.2019.00068.

Carreras Guzman, N.H., Wied, M., Kozine, I., Lundteigen, M.A., 2020. Conceptualizing the key features of cyber-physical systems in a multi-layered representation for safety and security analysis. Syst. Eng. 23 (2), 189–210. https://doi.org/10.1002/sys.v23.210.1002/sys.21509.

Carreras Guzman, N.H., Zhang, J., Xie, J., Glomsrud, I.A., 2021. A Comparative Study of STPA-Extension and the UFoI-E Method for Safety and Security Co-analysis. Reliab. Eng. Syst. Saf. 211, 107633.

Centre for Cyber Security, 2019. Threat Assessment: The Cyber Threat Against Denmark 2019. Copenhagen. Retrieved from https://fe-ddis.dk/cfcs/Pages/cfcs.aspx.

Checkoway, S., Mccoy, D., Kantor, B., Anderson, D., Shacham, H., Savage, S., Kohno, T., 2011. Comprehensive Experimental Analyses of Automotive Attack Surfaces. In: USENIX Security Symposyum. San Francisco. Retrieved from http://www.usenix.org/events/security/tech/full_papers/Checkoway.pdf.

Chockalingam, S., Hadziosmanovic, D., Pieters, W., Teixeira, A., van Gelder, P., 2013. Integrated Safety and Security Risk Assessment Methods: A Survey of Key Characteristics and Applications. In: Critical Information Infrastructures Security. 8th International Workshop, CRITIS 2013. Revised Selected Papers: LNCS 8328 (Vol. 8328). Doi: 10.1007/978-3-319-03964-0.

de Ruijter, A., Guldenmund, F., 2016. The bowtie method: A review. Saf. Sci. 88, 211–218. https://doi.org/10.1016/j.ssci.2016.03.001.

DHS, NCCIC, & ICS-CERT. (2016). Recommended Practice: Improving Industrial Control System Cybersecurity with Defense-in-Depth Strategies.

DNV GL, 2016. Cyber security resilience management for ships and mobile offshore units in operation. Retrieved October 30, 2018, from http://www.gard.no/Content/21865536/DNVGL-RP-0496.pdf.

DRAGOS, 2017. CRASHOVERRIDE: Analysis of the Threat to Electric Grid Operations. Retrieved October 28, 2019, from https://dragos.com/wp-content/uploads/CrashOverride-01.pdf.

Duijm, N.J., 2009. Safety-barrier diagrams as a safety management tool. Reliab. Eng. Syst. Saf. 94 (2), 332–341. https://doi.org/10.1016/j.ress.2008.03.031.

Dunjó, J., Fthenakis, V., Vílchez, J.A., Arnaldos, J., 2010. Hazard and operability (HAZOP) analysis. A literature review. J. Hazard. Mater. 173 (1–3), 19–32. https://doi.org/10.1016/j.jhazmat.2009.08.076.

Dzung, D., Naedele, M., Von Hoff, T.P., Crevatin, M., 2005. Security for industrial communication systems. Proc. IEEE 93 (6), 1152–1177. https://doi.org/10.1109/JPROC.2005.849714.

E-ISAC, 2016. Analysis of the Cyber Attack on the Ukrainian Power Grid: Defense USe Case. Doi: 10.1159/000329982.

Ericson, C.A., 2005. Hazard Analysis Techniques for System Safety. John Wiley and Sons, Inc. Doi: 10.1002/0471739421.

Fridman, L., 2019. Self-Driving Cars: State of the Art (2019). Retrieved February 5, 2019, from https://deeplearning.mit.edu.

Friedberg, I., McLaughlin, K., Smith, P., Laverty, D., Sezer, S., 2017. STPA-SafeSec: Safety and security analysis for cyber-physical systems. J. Inform. Sec. Appl. 34, 183–196. https://doi.org/10.1016/j.jisa.2016.05.008.

Ge, X., Yang, F., Han, Q.L., 2017. Distributed networked control systems: A brief overview. Inf. Sci. 380, 117–131. https://doi.org/10.1016/j.ins.2015.07.047.

Gill, H., 2008. From Vision to Reality: Cyber-Physical Systems (ppt presentation). National Science Foundation, 1–29.

Hagen, I.B., Kufoalor, D.K.M., Brekke, E.F., Johansen, T.A., 2018. MPC-based Collision Avoidance Strategy for Existing Marine Vessel Guidance Systems. In: Proc. IEEE International Conference on Robotics & Automation (ICRA), pp. 7618–7623.

Harari, Y.N., 2015. Sapiens: A Brief History of Humankind (1st ed.). Harper.

He, H., Yan, J., 2016. Cyber-physical attacks and defences in the smart grid: a survey. IET Cyber-Phys. Syst.: Theor. Appl. 1 (1), 13–27. https://doi.org/10.1049/cps2.v1.110.1049/iet-cps.2016.0019.

Hovden, J., Albrechtsen, E., Herrera, I.A., 2010. Is there a need for new theories, models and approaches to occupational accident prevention? Saf. Sci. 48 (8), 950–956. https://doi.org/10.1016/j.ssci.2009.06.002.

Humayed, A., Lin, J., Li, F., Luo, B., 2017. Cyber-Physical Systems Security - A Survey. IEEE Internet Things J. 4 (6), 1802–1831. https://doi.org/10.1109/JIOT.2017.2703172.

Hutchins, E., Cloppert, M., Amin, R., 2011. Intelligence-driven computer network defense informed by analysis of adversary campaigns and intrusion kill chains. Leading Issues Inform. Warfare Sec. Res. 1.

IEC/TR 62061, 2010. Guidance on the application of ISO 13849-1 and IEC 62061 in the design of safety-related control systems for machinery.

IEC/TR 63069, 2019. Industrial-process measurement, control and automation - Framework for functional safety and security.

IEC/TS 62443-1-1, 2009. Industrial communication networks. Network and system security. Part 1-1: Terminology, concepts and models.

IEC 61508-1, 2010. Functional safety of electrical/electronic/programmable electronic safety-related systems - Part 1: General requirements.

IEC 61508-4, 2010. Functional safety of electrical/electronic/programmable electronic safety-related systems - Part 4: Definitions and abbreviations.

IEC 61511-1, 2016. Functional safety – Safety instrumented systems for the process industry sector – Part 1: Framework, definitions, system, hardware and application programming Requirements.

IEC 62443-2-1, 2010. Industrial communication networks - Network and system security - Part 2-1: Establishing an industrial automation and control system security program.

ISO/IEC 27000, 2017. Information technology - Security techniques - Information security management systems - Overview and vocabulary.

ISO/PAS 21448, 2019. Road vehicles - Safety of the intended functionality.

ISO/TR 22100-4, 2018. Safety of machinery - Relationship with ISO 12100 - Part 4: Guidance to machinery manufacturers for consideration of related IT-security (cyber security) aspects.

ISO 26262-1, 2011. Road vehicles - Functional safety - Part 1: Vocabulary.

Jazdi, N., 2014. Cyber physical systems in the context of Industry 4.0. In: Proceedings of 2014 IEEE International Conference on Automation, Quality and Testing, Robotics, AQTR 2014, 1–4. Doi: 10.1109/AQTR.2014.6857843.

Khaitan, S.K., McCalley, J.D., 2015. Design techniques and applications of cyberphysical systems: A survey. IEEE Syst. J. 9 (2), 350–365. https://doi.org/10.1109/JSYST.2014.2322503.

Khan, F.I., Abbasi, S.A., 1999. Major accidents in process industries and an analysis of causes and consequences. J. Loss Prev. Process Ind. 12 (5), 361–378. https://doi.org/10.1016/S0950-4230(98)00062-X.

Koscher, K., Czeskis, A., Roesner, F., Patel, S., Kohno, T., Checkoway, S., Savage, S., 2010. Experimental security analysis of a modern automobile. Proceedings - IEEE Symposium on Security and Privacy, 447–462. Doi: 10.1109/SP.2010.34.

Kriaa, S., 2016. Joint safety and security modeling for risk assessment in cyber physical systems. Université Paris-Saclay.

Kriaa, S., Bouissou, M., Laarouchi, Y., 2019. A new safety and security risk analysis framework for industrial control systems. Proc. Instit. Mech. Eng., Part O: J. Risk Reliab. 233(2), 151–174. Doi: 10.1177/1748006X18765885.

Kriaa, S., Bouissou, M., Piètre-Cambacédès, L., 2012. Modeling the Stuxnet attack with BDMP: Towards more formal risk assessments. In: 7th International Conference on Risks and Security of Internet and Systems, CRiSIS 2012. Doi: 10.1109/CRISIS.2012.6378942.

Kriaa, S., Pietre-Cambacedes, L., Bouissou, M., Halgand, Y., 2015. A survey of approaches combining safety and security for industrial control systems. Reliab. Eng. Syst. Saf. 139, 156–178. https://doi.org/10.1016/j.ress.2015.02.008.

Kriaa, S., Raspotnig, C., Bouissou, M., Piètre-Cambacedes, L., Karpati, P., Halgand, Y., Katta, V., 2013. Comparing Two Approaches to Safety and Security Modelling : BDMP Technique and CHASSIS Method System architecture used for the case study. In: Proceedings of the Enlarged Halden Programme Group Meeting. Storefjell, Norway: OECD Halden Reactor Project.

Langner, R., 2011. Stuxnet: Dissecting a cyberwarfare weapon. IEEE Secur. Priv. 9 (3), 49–51. https://doi.org/10.1109/MSP.2011.67.

Lee, E.A., 2006. Cyber-Physical Systems - Are Computing Foundations Adequate? In: NSF Workshop On Cyber-Physical Systems: Research Motivation, Techniques and Roadmap. Austin, TX, pp. 1–9. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.84.8011&rep=rep1&type=pdf.

Lee, E.A., 2008. Cyber Physical Systems: Design Challenges. In: 11th IEEE Int. Symp. on Object and Component-Oriented Real-Time Distributed Computing, 363–369. Doi: 10.1109/ISORC.2008.25.

Lee, E.A., Seshia, S.A., 2017. Introduction to Embedded Systems: A Cyber-Physical Systems Approach. The MIT Press. Doi: 10.1016/B978-0-12-386886-2.00001-1.

Lee, R.M., Assante, M.J., Conway, T., 2014a. German Steel Mill Cyber Attack. Industrial Control Systems. Retrieved from http://ics3.sans.org/media/ICS-CPPE-case-Study-2-German-Steelworks_Facility.pdf.

Lee, R.M., Assante, M.J., Conway, T., 2014b. Media report of the Baku-Tbilisi-Ceyhan (BTC) pipeline Cyber Attack. Retrieved from https://ics.sans.org/media/Media-report-of-the-BTC-pipeline-Cyber-Attack.pdf.

Leveson, N.G., 2004. A new accident model for engineering safety systems. Saf. Sci. 42 (4), 237–270.

Leveson, N.G., 2011. Engineering a safer world: systems thinking applied to safety. The MIT Press.

Leveson, N.G., Thomas, J.P., 2018. STPA Handbook. Doi: 10.2143/JECS.64.3.2961411.

Lund, M.S., Solhaug, B., Stølen, K., 2011. Model-Driven Risk Analysis: The CORAS Approach. Springer, Berlin, Heidelberg. Doi: 10.1007/978-3-642-12323-8.

Lundteigen, M.A., Gran, B.A., 2019. The need of improved methods to handle functional safety and cybersecurity in industrial control and safety systems. In: OECD Halden Programme Meeting.

Lyu, X., Ding, Y., Yang, S.-H., 2019. Safety and security risk assessment in cyber-physical systems. IET Cyber-Phys. Syst.: Theor. Appl. 4 (3), 221–232. https://doi.org/10.1049/cps2.v4.310.1049/iet-cps.2018.5068.

Marwedel, P., 2011. Embedded System Design - Embedded Systems Foundations of Cyber-Physical Systems. Doi: 10.1007/978-94-007-0257-8.

Michniewicz, J., Reinhart, G., 2014. Cyber-physical Robotics – Automated Analysis, Programming and Configuration of Robot Cells based on Cyber-physical-systems. Procedia Technol. 15, 566–575. https://doi.org/10.1016/j.protcy.2014.09.017.

Miller, C., Valasek, C., 2015. Remote Exploitation of an Unaltered Passenger Vehicle. Defcon 23, 2015, 1–91. Retrieved from http://illmatics.com/Remote Car Hacking.pdf.

Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., Ueda, K., 2016. Cyber-physical systems in manufacturing. CIRP Ann. – Manuf. Technol. 65 (2), 621–641. https://doi.org/10.1016/j.cirp.2016.06.005.

Monostori, L., 2014. Cyber-physical production systems: Roots, expectations and R&D challenges. Procedia CIRP 17, 9–13. https://doi.org/10.1016/j.procir.2014.03.115.

Nai Fovino, I., Masera, M., De Cian, A., 2009. Integrating cyber attacks within fault trees. Reliab. Eng. Syst. Saf. 94 (9), 1394–1402. https://doi.org/10.1016/j.ress.2009.02.020.

National Highway Traffic Safety Administration, 2016. Cybersecurity Best Practices for Modern Vehicles. US Department of Transportation, 22p. Retrieved from http://www.nhtsa.gov/staticfiles/nvs/pdf/812333_CybersecurityForModernVehicles.pdf%0Ahttps://trid.trb.org/view/1428799.

Park, K.J., Zheng, R., Liu, X., 2012. Cyber-physical systems: Milestones and research challenges. Comput. Commun. 36 (1), 1–7. https://doi.org/10.1016/j.comcom.2012.09.006.

Paul, S., 2015. On the meaning of security for safety (S4S). In: Safety and Security Engineering VI, pp. 379–389. Doi: 10.2495/safe150321.

Paul, S., Brunel, J., Rioux, L., Vallée, F., de Oliveira, J., Gailliard, G., Chemouil, D., 2016. Recommendations for security and safety co-engineering (Release No. 3). MeRgE ITEA2 Project.

Pietre-Cambacedes, L., Chaudet, C., 2010. The SEMA referential framework: Avoiding ambiguities in the terms “security” and “safety”. Int. J. Crit. Infrastruct. Prot. 3 (2), 55–66. https://doi.org/10.1016/j.ijcip.2010.06.003.

Pinto, A. Di, Dragoni, Y., Carcano, A., 2018. TRITON: The First ICS Cyber Attack on Safety Instrument Systems. In: Black Hat USA 2018. Retrieved from https://www.nozominetworks.com/downloads/US/Nozomi-Networks-TRITON-The-First-SIS-Cyberattack.pdf.

Plioutsias, A., Karanikas, N., Chatzimihailidou, M.M., 2018. Hazard analysis and safety requirements for small drone operations: to what extent do popular drones embed safety? Risk Anal. 38 (3), 562–584. https://doi.org/10.1111/risa:2018.38.issue-310.1111/risa:12867.

Protti, M., Barzan, R., 2007. UAV Autonomy – Which level is desirable ? – Which level is acceptable ? Alenia Aeronautica Viewpoint. In: Platform Innovations and System Integration for Unmanned Air, Land and Sea Vehicles (AVT-SCI Joint Symposium), 1–12. Retrieved from http://www.rto.nato.int/abstracts.asp.

Rajkumar, R., de Niz, D., Klein, M., 2017. Cyber-Physical Systems, 1st ed. Addison-Wesley Professional.

Rajkumar, R., Lee, I.L.I., Sha, L.S.L., Stankovic, J., 2010. Cyber-physical systems: The next computing revolution. In: Design Automation Conference (DAC), 2010 47th ACM/IEEE, 0–5. Doi: 10.1145/1837274.1837461.

Rasmussen, B., Grønberg, C.D., 1997. Accidents and risk control. J. Loss Prev. Process Ind. 10 (5–6), 325–332. https://doi.org/10.1016/S0950-4230(97)00022-3.

Raspotnig, C., 2014. Requirements for safe and secure information systems. University of Bergen.

Raspotnig, C., Katta, V., Karpati, P., Opdahl, A.L., 2013. Enhancing CHASSIS: A method for combining safety and security. In: Proceedings - 2013 International Conference on Availability, Reliability and Security, ARES 2013, (1751), 766–773. Doi: 10.1109/ARES.2013.102.

Raspotnig, C., Opdahl, A., 2013. Comparing risk identification techniques for safety and security requirements. J. Syst. Softw. 86 (4), 1124–1151. https://doi.org/10.1016/j.jss.2012.12.002.

Rausand, M., 2011. Risk Assessment: Theory, Methods, and Applications. John Wiley & Sons.

Reason, J. (Ed.), 1990. Human Error. Cambridge University Press.

Schmittner, C., Ma, Z., Puschner, P., 2016. Limitation and Improvement of STPA-Sec for Safety and Security Co-analysis. In: Skavhaug, A., Guiochet, J., Schoitsch, E., Bitsch, F. (Eds.), Computer Safety, Reliability, and Security: SAFECOMP 2016 Workshops ASSURE, DECSoS, SASSUR, and TIPS. Trondheim, pp. 195–209. Retrieved from https://doi.org/10.1007/978-3-319-45480-1.

Schmittner, C., Ma, Z., Schoitsch, E., Gruber, T., 2015. A case study of FMVEA and CHASSIS as safety and security co-analysis method for automotive cyber-physical systems. In: CPSS 2015 - Proceedings of the 1st ACM Workshop on Cyber-Physical System Security, Part of ASIACCS 2015, 69–80. Doi: 10.1145/2732198.2732204.

Schneier, B., 1999. Attack Trees. Retrieved September 11, 2019, from https://www.schneier.com/academic/archives/1999/12/attack_trees.html.

Simensen, J.E., Sarshar, S., Hauge, A.A., Olsen, S.A., Sechi, F., Jørgensen, P.-A., 2019. HWR-1244 Test Enclave - Technical Specification Documentation. (For use within the Halden Project Member Organizations only). OECD Halden Reactor Project.

Slay, J., Miller, M., 2007. Lessons Learned from the Maroochy Water Breach. In: Goetz, E., Shenoi, S. (Eds.), IFIP International Federation for Information Processing, Vol. 253. Springler, Boston, pp. 73–82. https://doi.org/10.1007/978-0-387-75462-8.

Stouffer, K., Pillitteri, V., Lightman, S., Abrams, M., Hahn, A., 2015. Guide to Industrial Control Systems (ICS) Security. National Institute of Standards and Technology (NIST). Doi: 10.6028/NIST.SP.800-82r2.

Taylor, J.R., 2017. Automated HAZOP revisited. Process Saf. Environ. Prot. 111, 635–651. https://doi.org/10.1016/j.psep.2017.07.023.

Tegmark, M., 2018. Life 3.0: Being Human in the Age of Artificial Intelligence. Vintage.

Temple, W., Wu, Y., Chen, B., Kalbarczyk, Z., 2017. Reconciling systems-theoretic and component-centric methods for safety and security co-analysis. In: Tonetta, S., Schoitsch, E., Bitsch, F. (Eds.), Computer Safety, Reliability, and Security: SAFECOMP 2017 Workshops ASSURE, DECSoS, SASSUR, TELERISE, and TIPS. Trento, pp. 87–93. Doi: 10.1007/978-3-319-66284-8.

Torkildson, E.N., 2018. Empirical Studies of Safety and Security Co-analysis of Autonomous Systems. Norwegian University of Science and Technology (NTNU).

Weiss, J., 2010. Protecting Industrial Control Systems from Electronic Threats. Momentum Press.

Yampolskiy, M., Horvath, P., Koutsoukos, X.D., Xue, Y., Sztipanovits, J., 2012. Systematic analysis of cyber-attacks on CPS-evaluating applicability of DFD-based approach. In: Proceedings - 2012 5th International Symposium on Resilient Control Systems, ISRCS 2012, 55–62. Doi: 10.1109/ISRCS.2012.6309293.

Young, W., Leveson, N.G., 2013. Systems thinking for safety and security. In: Proceedings of the 2013 Annual Computer Security Applications Conference (ACSAC), 1–8. Doi: 10.1145/2523649.2530277.

Zeller, M., 2011. Myth or reality - Does the Aurora vulnerability pose a risk to my generator? In: 2011 64th Annual Conference for Protective Relay Engineers. IEEE, pp. 130–136. Doi: 10.1109/CPRE.2011.6035612.

Zio, E., 2018. The future of risk assessment. Reliab. Eng. Syst. Saf. 177, 176–190. https://doi.org/10.1016/j.ress.2018.04.020.