TU Delft

Delft University of Technology

Recent Cyber-Physical-System developments and their safety & security management risk factors

Lindhout, Paul; Reniers, Genserik

DOI
10.59490/pss.1.2025.8097

Publication date
2025

Document Version
Final published version

Published in
Journal of Progress in Safety & Security

Citation (APA)
Lindhout, P., & Reniers, G. (2025). Recent Cyber-Physical-System developments and their safety & security management risk factors. *Journal of Progress in Safety & Security*, *1*. https://doi.org/10.59490/pss.1.2025.8097

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Review article

TU Delft OPEN Publishing

Recent Cyber-Physical-System developments and their safety & security management risk factors

Paul Lindhout [1,2], Genserik L.L.M.E. Reniers [1,3,4,*]

[1]Faculty of Applied Economic Sciences and Engineering Mgmt (ENM), University of Antwerp, Belgium
[2]Dept. of Care Ethics, University for Humanistic Studies (UvH), The Netherlands
[3]Center for Corporate Sustainability (CEDON), KU Leuven, Belgium
[4]Faculty TPM-VTI (S3G), Delft University of Technology, The Netherlands

*Corresponding author: genserik.reniers@uantwerpen.be

Abstract: Introduction: Recent developments in a wide variety of cyber-physical system (CPS) designs have resulted in many safety and security concerns. These are caused by increasing complexity, connectivity to the outside world, and a variety of component types, all contributing to increased system vulnerability. Simultaneously, the use of digital models, big data analytics, remote computing, online services, and machine learning extends manufacturing systems far beyond their physical and visible boundaries.
Focusing on industry and healthcare settings, this study explores the challenges faced by safety and security management originating from these developments.
Method: The scoping review method was used to gather literature on safety and security concerns. These are grouped and linked to the CPS components. Subsequently, their impact is reviewed from the perspective of safety and security management systems.
Results: The findings indicate that safety and security management become entangled, many new risks need to be considered, and the availability of online digital real-time system components and employee skills are critical factors. Some of the sensors and safety metrics software must now be classified as safety-critical because robots and people have started to share the same workspace. The rapid developments and many unresolved issues on design, ethics, data and software quality, external supervision, standards and guidance, legal frameworks, and online system components abroad urgently require attention.
In the discussion, several recommendations are made regarding how safety and security management can address these challenges.
Conclusions: While robots and people increasingly mingle on shop floors and in care situations, safety management is lagging behind current CPS developments. New critical safety aspects require the attention of safety and security management.
Practical applications: These findings in safety and security management systems will contribute to avoiding harm to people and the environment in companies operating CPS.

ISSN: 3050-4570
Vol. 1, 2025
DOI: 10.59490/pss.1.2025.8097

©2025 Authors. Published by TU Delft OPEN Publishing on behalf of the authors. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

One sentence summary: This study presents a structured way to deal with safety management of Cyber-Physical-Systems.

Keywords: Cyber Physical System, Safety and Security management system, concerns, issues, vulnerability.

Publishing history: Submitted: 19 February 2024; Revised: 7 April 2024; Accepted: 31 March 2025; Published: 21 July 2025

Cite as: Lindhout, P., Reniers, G.L.L.M.E. (2025) Recent Cyber-Physical-System developments and their safety & security management risk factors. *Journal of Progress in Safety & Security, 1*. https://doi.org/10.59490/pss.1.2025.8097

©2025 Authors. Published by TU Delft OPEN Publishing on behalf of the authors. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence.

1 Introduction

Although cyber-physical systems (CPS) do not have a single undebated definition (Putnik et al., 2019), they have common properties. A recent CPS consisting of a physical process running in an electromechanical installation is controlled by digital computers, software, and data in a system with decision-making, self-adjusting, predicting, and learning functionality, which is situated for a part locally on the premises, monitoring its direct environment via sensors, interacting with humans, and being connected to remote locations and services on the Internet.

CPS designs are a subject of concern, both from a safety point of view and from a security point of view (Chaminda, 2021; Reniers & Amyotte, 2012). These two disciplines meet in the underground caverns of design gaps, programming errors, and unauthorized access in a dynamic and complex setting. The CPS design is "*not yet fully mature*" (Vogt, 2021, p1085).

Nonetheless, a growing number of CPS variant designs are being developed. These CPS designs are increasingly complex (Van Acker, 2020; Haidegger et al., 2020), increasingly connected to the outside world (Gromov, 1995; Peserico et al., 2021; Purcell, 2012; Cho, 2020) and using more information (Malhotra et al., 2021), software (Wirth, 1995), digital models (Wang et al., 2021) and data (Gray et al., 2005) than ever before. Data are generated in real-time within the organization (Hopcraft et al., 2021; Khalid et al., 2017; Fletcher & Webb, 2017) and outside the organization (Purcell, 2012; Asaad et al., 2021).

Several new technologies are used to support such systems, such as the Internet of Things (IoT) (Purcell, 2012), remote computing (McKee et al., 2017; Leiber, 2017; Girs et al., 2020), artificial intelligence (AI) (Kline, 2011; Deschacht, 2021) and machine learning (ML) (Bharadwaj et al., 2021).

A generic configuration of a CPS, inspired by the works of Jiang et al. (2019), Mezzanotte (2019), and Chehri et al. (2021), is shown in figure 1.

At first glance, a CPS is designed and maintained according to various standards and is intended to perform an operational activity, for example, robotic manufacturing of a product, running a chemical process, or performing a routine surgery, all of which require a flow of materials and/or services. It consists of interconnected components that can be physically located both inside and remotely. These components include hardware, software, and data. This system is supported by utilities, people, and digital networks. Its capabilities are determined by the dedicated design and use of standards. Its performance is expressed as product or service output. External supervision (governance) must ensure ethical conduct and compliance with regulatory requirements.

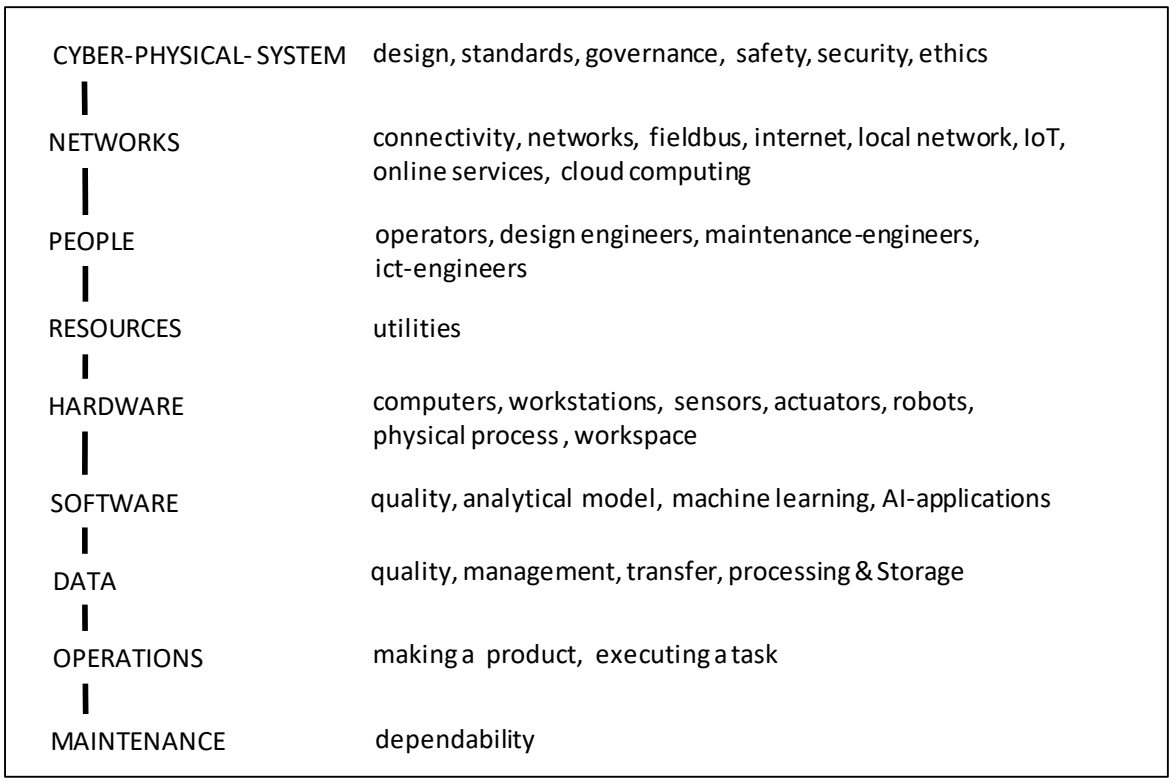

Figure 1 - Generic Cyber-Physical-System component structure and risk factors

2 Problem Definition

These systems present new and increasingly more safety and security management challenges to organizations and the people involved in CPS design, development, and operations.

Within these systems, human and robotic workplaces are merged (Patriarca et al., 2021) where they have thus far been separated for safety reasons. Within such systems, automatic safety devices with millisecond response times can be employed, making human intervention impossible (Peserico et al., 2021). CPS safety is becoming a major concern and requires improvement (Arden et al., 2010; Vogt, 2021). Both connectedness with the outside world and the integrated work environment of a CPS via the Internet also make security a major concern (Reniers & Amyotte, 2012; Olowononi et al., 2020). The research questions in this study were as follows:

> *Which challenges to health, safety and security management are emerging due to the development of an increasingly wider variety of Cyber-Physical-Systems?*

The authors explore current CPS safety and security issues and how these can be addressed in safety and security management systems (SSMS) in industry and healthcare.

3 Materials and Methods

To this end, the authors conducted a literature study following the scoping review approach (Smith et al., 2015), consisting of a preparatory search, followed by multiple successive searches with progressively refined combinations of search terms, and additional searches to find background or detailed information. We developed a set of initial search terms in three steps.

1-The choice of the primary key search terms safety, security, management, industry and health care follows from the scope of this study.
2-The search terms occupational, risk inventory, hazard, problem, concern, issue, decision, trend and future development are important due to the nature of the problem at hand.
3-Specific technical search terms are drawn from sources identified in preparatory searches: big data, artificial intelligence, cyber-physical systems, cyber-socio-technical systems, software, algorithms, and other terms as encountered.

During multiple combination searches, other relevant search terms may be encountered, which are considered in further combination searches. Several additional studies have explored specific areas (Byrne, 2016; Byrne, Daykin & Coad, 2016). With a few notable exceptions, mainly related to historical developments, the time period for searching is limited to after 2015, in line with the emergence of the relevant literature (Forcina & Falcone, 2021; Dey & Lee, 2021; Jafari et al., 2022). Potentially relevant sources were gathered and selected based on relevance, first on title, abstract, and keywords, and then on full text content.

The research question and sub-questions cannot be properly addressed based only on scientific literature. Several non-scientific information sources must be considered. We distinguish between 1) primary peer-reviewed scientific sources, 2) secondary sources generated by universities and international public or governmental organizations, and 3) tertiary ‘grey’ sources generated by others (Wessels, 1997; Pandita & Singh, 2011; Cronin et al., 2008).

Due to the general nature of the subject, several non-specialized databases were used for the literature search: Google Scholar and its associated proprietary databases, Academia, and Research Gate. Duplicate sources, sources that do not address safety and security aspects, sources that do not have a bearing on industry or healthcare, and studies that are not available in English were excluded. This study yielded 190 relevant sources. Table 1 summarizes the literature search and selection processes.

The findings were derived from these sources via simple text analysis and meta-synthesis (Cronin et al., 2008; Noah, 2017).

Date	Description	Search results Returned [hits]	[cut-off]	Sources admitted per type [Number of sources] 1)	2)	3)	Total
Preparatory searches							
28-3-2022	1 background info	n/a	n/a	-	-	-	-
30-3-2022	2 digital technology	n/a	n/a	14	3	11	28
2-4-2022	3 method	n/a	n/a	7	0	0	7
7-4-2022	4 cyber socio technical model	14.4K	1st 100	8	0	0	8
31-3-2022	5 internet of things growth	17.6K	1st 100	5	0	1	6
n/a	6 references search	n/a	n/a	3	0	0	3
n/a	7 sources available	n/a	n/a	5	0	0	5
Combination searches							
25-4-2022	8 hazards	10	n/a	4	1	0	5
27-4-2022	9 health care,1st attempt	60	n/a	5	0	0	5
28-4-2022	9 health care,2nd attempt	44	n/a	1	0	0	1
29-4-2022	10 industry, 1st attempt	1	n/a	1	0	0	1
29-4-2022	10 industry, 2nd attempt	1	n/a	1	0	0	1
29-4-2022	10 industry, 3rd attempt	18.7K	1st 100	13	0	0	13
4-5-2022	11 safety mgmt system	35	n/a	5	0	0	5
Additional searches							
24-5-2022	12 reference listings	46	n/a	22	0	6	28
25-5-2022	13 related standards	57	n/a	0	43	0	43
29-5-2022	14A safety decision making	3	n/a	2	0	0	2
29-5-2022	14B safety decision making	16.9K	1st 100	11	0	0	11
9-7-2022	15 ethical design for AI	8	n/a	2	0	4	6
12-7-2022	16 decision making RG/GS	12	n/a	12	0	0	12
				___+	___+	___+	___+
			Total	121	47	22	190

Table 1 – The literature search and selection process; type 1) primary scientific sources, type 2) secondary sources generated by universities and international organizations and type 3) tertiary 'grey' sources.

4 Results

Although systems with new architectures are being developed, they incorporate a growing number of different component types. Terminology is still under development, as indicated by the growing number of names and abbreviations for systems found in the literature.

Cyber-Physical Systems (CPS) is a system with interconnected hardware, software, intelligent functionality, real-world sensing, and physical and electromechanical components (Mezzanotte, 2019; Yilma et al., 2021; Vierendeels et al., 2018);

Human-Cyber-Physical Systems (HCPS) – a CPS supported and integrated with human knowledge and expertise using AI modelling techniques (Bocklisch et al., 2022);

Cyber-Physical-Social System (CPSS) – a CPS with an added-on social dimension, with increased use of smart devices, sensors, and a direct link with operators/users (Cabour et al., 2021; Yilma et al., 2021)

Cyber-Physical System of Systems (CPSoS): a System of Systems, multiple cooperating or integrated CPS installations (Gharib et al., 2021; Matta et al., 2021);

Cyber-Socio-Technical Systems (CSTS) – a CPS cooperating within an interconnected group of multiple other cyber technical systems, physical devices, and humans (Patriarca et al., 2021);

Collaborative Cyber-Physical System (CCPS) – a CPS with provisions for safe human – machine interaction on the shop floor (Khalid et al., 2018);

Collaborative Robotic Cyber-Physical System (CRCPS) – a CPS with safe and secure human-robot collaboration (Khalid et al., 2018);

Medical Cyber-Physical Systems (MCPS) are CPS built for medical applications (Liu, 2022).

Although different names are in use for Cyber Physical Systems (CPS) today, the authors refer to all of them with "CPS." In addition to the description in the introduction, in this study, a CPS is considered as physically operating in the presence of people and as having some form of external online support, being automated, having some level of autonomy, being enabled for a part to sense its environment, making use of artificial intelligence, and having machine learning capability. Most likely, a CPS also extends beyond the physical perimeter of a factory production site, an operating theatre in a hospital, or the bridge of a cruise liner and is connected with virtual online and real-time components in the cyber world, physically located elsewhere in the physical world, and with people involved in its operations, either inside or outside the organization.

The results of the literature search are presented in the following two sections, first sketching the background and then indicating the impact on safety and security management. In the first of these two sections, CPS design, development, and operational aspects, which have an impact on safety and security, are gathered. The second section CPS CPS-related concerns and issues in the SSMS section.

4.1 CPS design, development and operational aspects

The current 4th industrial revolution (Schwab, 2015) introduced new technologies with a profound effect on CPS safety and security risk profiles. The concerns and issues reported in the literature are reviewed here, following the CPS component structure shown in figure 1.

4.1.1 CYBER-PHYSICAL-SYSTEM

4.1.1.1 DESIGN and STANDARDS management risk factors

Adding complexity generally means new or increased vulnerability to error (Peserico et al., 2021). In addition, the heterogeneity of components and their interoperability (Malhotra et al., 2021; Angelopoulos et al., 2019) can exacerbate this problem.

Standards relevant to CPS design follow rapid technological development, rather than guiding and structuring it. This leads to several safety and security problems. First, not all CPS can be appraised against a law, directive, guidance, protocol, or other documented standard (Begic & Galic, 2021). There are many areas in CPS design where a yardstick is missing, leading to CPS designs not being subject to a well-defined set of safe operating values, standards, classifications, or well-defined criteria for acceptability (Magrabi et al., 2019; Pereira & Thomas, 2020; Fosch-Villaronga & Mahler, 2021). The absence of a safe bus communication protocol (Huang et al., 2019), and the few standards that address human-robot interactions (Olszewska et al., 2020) compound this problem. There is no guidance for the design of a suitable sensor arrangement, which is essential for enabling a CPS to observe its environment and, more importantly, human presence and behavior (Khalid et al., 2017). Apart from the general product safety requirements (EC 2023/988), there are 42 standards applicable to CPS safety and security (See appendix-1). These were identified from literature sources included in this study (Mezzanotte, 2019; Coscia et al., 2021; Fosch-Villaronga & Mahler, 2021; Haidegger et al., 2020; Pereira & Thomas, 2020; Broum & Šimon, 2020; EC, 2020a; Gualtieri et al., 2018; Medina et al., 2019). These standards are not further discussed here but they are included with full reference in Appendix-1 for convenience (ANSI/RIA R15.06; EN 62304:2015; ETSI TR 103375; IEC 60601:2017; IEC 61508:2016; IEC 62046:2018; IEC 62541; IEEE 1872:2015; IEEE P1872.2; IEEE P7001:2017; IEEE P7000:2017; ISO 10075; ISO 10218-1,-2:2011; ISO 12100:2021; ISO 13482:2014; ISO 13485:2003; ISO 13854:2017; ISO 13855:2010; ISO 13857:2008; ISO 14001; ISO 14118:2015; ISO 14120:2015; ISO 26000; ISO 26262:2018-12; ISO 26800; ISO 31000; ISO 45001; ISO 6385; ISO 8373:2012; ISO 9000; ISO 9241; ISO/CD 8373:2019; ISO/DTR 23482-1; ISO/IEC 18033; ISO/IEC 27001; ISO/IEC 27002; ISO/IEC 27040; ISO 27701:2019; ISO/IEC 29100; ISO/PRF TR 23482-2; ISO/TR 16982; ISO/TS 13849:2018; ISO/TS 15066:2016; SAE ARP4754A:2010). Although this is not an exhaustive list because new standards and versions continue to appear, and a lack of standards exists in several areas, this large number of standards underlines the magnitude of the challenge for design engineers when a CPS is newly developed in their organization.

4.1.1.2 EXTERNAL SUPERVISION (Governance) risk factors

Accountability and liability are concerns (Coscia et al., 2021; Tschider, 2018; Magrabi et al., 2019; Khalid et al., 2018) because an organization exploiting CPS involves a growing number of stakeholders (Fosch-Villaronga & Mahler, 2021). In addition, societal and environmental impacts (Coscia et al., 2021; Amodei et al., 2016), data management (Coscia et al., 2021; Koene et al., 2018) and ethical aspects (Garibaldo & Rebecchi, 2018) pose challenges to those involved in the external supervision (governance) of such organizations.

4.1.1.3 SAFETY management risk factors

There is a lack of consistent safety and security guidelines (Johnsen et al., 2009; Haidegger et al., 2020; Leong, 2018) and knowledge sharing (Stefana & Paltrinieri, 2021). Standards and guidance regarding safety are also lagging behind, incomplete, or inconsistent (Fosch-Villaronga & Mahler, 2021; McKee et al., 2017; Begic & Galic, 2021; Olszewska et al., 2020; Leong, 2018). In some areas, there are overlapping or conflicting standards (Coscia et al., 2021).

A lack of standards is specifically felt in software architecture (Mezzanotte, 2019), system certification (Assis-Dornelles et al., 2022; Johnsen et al., 2009), and ethics and safety guidelines (Haidegger et al., 2020; Koene et al., 2018; Vinuesa et al., 2018; Olszewska et al., 2020). In healthcare, two main application areas for CPS assistance require attention when it comes to safety. The first concerns the physical safety of the patient on the operating table when subjected to CPS-assisted surgery. The second concerns CPS offering social services in a care institution, such as serving food, distributing medicine, or providing robotic social contact and companionship.

Astonishingly, these 'care robots' are not classified in a clear way and standards do not address ensuring a safe human–robot interaction (Fosch-Villaronga & Mahler, 2021). In practice, a CPS workplace is safety-protected by 79% to 90.8% (Medina et al., 2019). If the system fails, it may harm people in areas where robot movement can cause impact or crush hazards.

4.1.1.4 SECURITY management risk factors

Cybersecurity is a major concern. Security breaches via unauthorized access by a hacker (Jiang et al., 2019), data damage, (non) intentional abuse, and reward hacking (Dey & Lee, 2021; Van Acker, 2020; Amodei et al., 2016; Sculley et al., 2015; Leslie, 2019; Zimmerman & Renaud, 2019; Google, 2019) have been reported in the literature. In addition, the absence of network barriers (Johnsen et al., 2009; Soldatos & Kyriazis, 2021) can lead to misleading or biased ML algorithms (Leslie, 2019) or to an "adversarial exploit" of vulnerabilities (Jiang et al., 2019).

4.1.1.5 ETHICS management risk factors

Ethics in relation to CPS are frequently addressed in the literature. More attention to ethics is required in CPS design and operation. Aspects mentioned are e.g. discrimination (Magrabi et al., 2019; Fosch-Villaronga & Mahler, 2021; Coscia et al., 2021; Tschider, 2018), human rights, diversity, fairness, disparate impact (Coscia et al., 2021; Amodei et al., 2016; Hamon et al., 2020), privacy or anonymity (Mezzanotte, 2019; Ogbuke et al., 2022; Marjumin et al., 2019; Malhotra et al., 2021; Tschider, 2018, Costantino et al., 2021; Van Acker, 2020; Coscia et al., 2021; Amodei et al., 2016; Liu, 2022; Fosch-Villaronga & Mahler, 2021; Furstenau et al., 2022; Olszewska et al., 2020; Leiber, 2017; Spiekermann & Winkler, 2020), dignity (Fosch-Villaronga & Mahler, 2021) and non-participation in decisions about data (Leiber, 2017; Fletcher & Webb, 2017). Other ethical aspects include proper handling of freedom of speech and mental health (Trentesaux & Karnouskos, 2019; Tschider, 2018; Laurent & Fabiano, 2022; Koene et al., 2018; Spiekermann & Winkler, 2020) are less frequently mentioned. The guiding role of IEEE must be mentioned here. The P7000 series of standards is under development, inspired by the vision laid down in their Global Initiative on Ethics and Ethically Aligned Design documents (IEEE, 2016, 2017).

4.1.2 NETWOKS

4.1.2.1 CONNECTIVITY and DYNAMICS management risk factors

Networks, Fieldbus, Internet, Local network, Internet of Things (IoT), and other interconnections, together with their mutual dynamic interactions, are major challenges when it comes to reliable operations. This makes a CPS vulnerable to unexpected interactions and influences from the outside.

Current systems show increased connectivity (Khalid et al., 2018; Olowononi et al., 2020; EC, 2020b; Costantino et al., 2021), and exposure to the outside world via an increased IoT "contact surface" (Pogliani et al., 2019), creating interdependency and interactivity (Ale et al., 2014) with others, in some cases artificial intelligence, automation, and autonomous systems (Stefana & Paltrinieri, 2021).

4.1.2.2 ONLINE SERVICES, CLOUD COMPUTING management risk factors

In a CPS, several online services can be used for input, for example, servers providing real-time weather data, fast calculation capacity, AI applications, and predicting simulation models. This raises issues such as delays, interruptions, errors, non-availability, and changes (Peserico et al., 2021).

4.1.3 PEOPLE

4.1.3.1 OPERATORS management risk factors

The non-existing, unprotected, limited, or poorly defined human role in a cyber-technical system (Peserico et al., 2021; Soldatos & Kyriazis, 2021; Spiekermann & Winkler, 2020) can endanger people during interaction with the system. This challenges design engineers to foresee such dangers for operators, maintenance engineers, and ICT engineers, for example, during system testing.

4.1.3.2 DESIGN ENGINEERS management risk factors

A CPS configuration may be subject to changes in tasks, configurations, or software, leading to dynamic risks (Mol, 2003). Meanwhile, microminiaturization continues (Moore, 1965; Roser & Ritchie, 2020) and future CPS complexity is likely to increase further.

4.1.3.3 MAINTENANCE ENGINEERS management risk factors

Maintenance engineers can be subjected to safety risks when working on a CPS during normal operation, testing, repair, or modification activities. Such risks are particularly important in a workspace shared between people, robots, sensors, and actuators (Patriarca et al., 2021).

Sudden movements of a robot or an actuator, for example, incurred by automatic response on a false alarm (Pereira & Thomas, 2020), by blocking a sensor (Liu, 2022), wrongly interpreted speech commands (Khalid et al., 2017), unintended triggering of an automatic safety device, unexpected – not foreseen – situations (Zheng et al., 2017), wrong decisions (See et al., 2017), improper patching (Cantelmi et al., 2021), or due to unavailability of safety measures such as restrained speed, can cause impact, crush danger, and fatalities (Khalid et al., 2017; See et al., 2017). Many other occupational hazards in the CPS environment have been identified by Costantino et al. (2021).

4.1.3.4 ICT ENGINEERS management risk factors

ICT engineers are involved in the development and maintenance of a CPS and can be exposed to dangers such as uncontrolled and potentially dangerous software updates (Spiekermann & Winkler, 2020; EC, 2020b), inadequate patches (Johnsen et al., 2009) or insufficient maintenance (Liu, 2022).

Poor diagnostics, that is, the CPS does not adequately detect its own failures well enough (McKee et al., 2017; Sculley et al., 2015; Vinuesa et al., 2018), can degrade maintenance to breakdown management, characterized by repair after a fault, loss of production, and downtime. These flaws might expose engineers to unpredictable CPS motion, impact and crushing hazards, and other dangers associated with wrongly handled dangerous chemicals, heat, or electricity.

The tasks that a CPS must perform are increasingly complex (Van Acker, 2020). The complexity of interactions between many system components, humans moving in and out of the joint working space (Yilma et al., 2021) and interacting AI applications (Soldatos & Kyriazis, 2021; Stefana & Paltrinieri, 2021; Liu, 2022) appear to be reaching beyond the ability of many people to understand what is going on. This lack of understanding might affect the workers' physical safety or lead to significant installation damage. Whether it is the limited capacity of the human brain (Benna & Fusi, 2016) or a challenge to the engineers and operators skills set, referred to as "data literacy" (Wang & Wang, 2021, p5), the ability to handle system complexity is becoming a new Human Resource management (HRM) requirement for people involved with a CPS.

4.1.4 RESOURCES

4.1.4.1 UTILITIES management risk factors

CPS resources can be disturbed at any moment. Examples of this are non-availability, time-out, back-up, and shut-down of utilities such as 220V power supply, internet, bus system, Wi-Fi, 4G/5G, and cloud service (Fosch-Villaronga & Mahler, 2021; Jing et al., 2014; Dey & Lee, 2021; Tschider, 2018; Furstenau et al., 2022; EC, 2020a). Because a CPS employs increasingly more physical or online service suppliers than before, their continuous and real-time supply of service may be subject to interruption as well (Fosch-Villaronga & Mahler, 2021; Mageto, 2021).

4.1.5 HARDWARE

4.1.5.1 COMPUTERS, WORKSTATIONS, ACTUATORS, ROBOTS, WORKSPACE, PHYSICAL PROCESS management risk factors

Although computers and workstations have their own occupational safety characteristics, they were not addressed in this study. Moving parts belonging to a CPS, such as actuators and robots, but also AGV's, CNC-lathes, 3D printers, and many other devices, can cause harm to people. This can have several different causes, such as software errors, malfunctions, or cyber-attacks (Peserico et al., 2021; Olowononi et al., 2020). The manufacturing process or other tasks performed by a CPS have their own risk inventories. These factors were not explored further in this study.

The human-machine relationship in a CPS setting makes it necessary from a workplace occupational safety point of view to consider both physical and mental safety aspects (Pena-Casas et al., 2018; Costantino et al., 2021).

Physical safety can differ between industrial production environments and medical or care settings. Workers in an industrial CPS setting can become victims of poor sensor-induced robot zone perimeter violations (Khalid et al., 2017) or CPS failure. Medical staff, care personnel, and patients can receive shocks or burns, for example, from wearable devices or medical devices (EC, 2020a; Broum & Šimon, 2020; Calvetti et al., 2020). Any hardware part can both fail instantaneously (Khalid et al., 2018) or show performance degradation, either by itself or by resource or connectivity failure (Bellini et al., 2021; Soldatos & Kyriazis, 2021; Bolbot et al., 2019; Hussain & Hong, 2020; Pereira & Thomas 2020; Liu, 2022). Safety-critical hardware may not meet SIL-3 level requirements or cannot be reliably coordinated (Peserico et al., 2021).

Workers and operators can feel fear, anxiety, insecurity, fatigue, or uncomfortable and face challenges in their mental health (Soldatos & Kyriazis, 2021; Moencks et al., 2022; Costantino et al., 2021; Van Acker, 2020; Medina et al., 2019; EC, 2020a; Spiekermann & Winkler, 2020; Fosch-Villaronga & Mahler, 2021).

Currently, there is no insight into the associated long-term effects on workers' health and safety when it comes to working in CPS settings (Moencks et al., 2022). At the same time, the potential for new causes of occupational diseases cannot be ruled out at this time (Costantino et al., 2021).

4.1.5.2 SENSORS management risk factors

Sensors are a category of CPS components that require particular attention. For example, a single broken proximity sensor can cause harm (Hussain & Hong, 2020). The CPS design and its sensor types and their arrangement determine how well the CPS can 'see' its dynamic environment and the humans working there (Coscia et al., 2021; Khalid et al., 2017; Soldatos & Kyriazis, 2021).

Especially poor image capture by sensors or cameras is unsafe when used in a CPS common robot and human joint workspace since the CPS simply does not 'see' the workers (Magrabi et al., 2019; Hussain & Hong, 2020). Robots in medical applications may not comply with medical-device standards (Fosch-Villaronga & Mahler, 2021). A sensor's field of view may be subject to limitations for safety reasons (Sun et al., 2020). In addition, sensor data rates can constrain their efficacy in monitoring the location of humans (Khalid et al., 2017).

4.1.6 SOFTWARE

4.1.6.1 QUALITY management risk factors

Software is a market-driven commodity that is fragmented. Its hierarchical and layered architecture may even be subjected to "competing . . . initiatives' (Mezzanotte, 2019, p14). A historically grown variety of programming languages are available to programmers. Incompatibility between programming languages may occur (Liu, 2022). Ongoing software explosions create an equally fast-growing need for programmers (Wirth, 1995). The risk of errors, flaws, glitches, and inadequacies (Bellini et al., 2021; Bolbot et al., 2019; Fosch-Villaronga & Mahler, 2021; Hussain & Hong, 2020) in software manifests itself in poor or improperly designed algorithms (O'Neil, 2017; Tramarin et al., 2019; Pereira & Thomas, 2020; Hamon et al., 2020; Tschider, 2018; Koene et al., 2018), poor models, poor predictions, and poor machine learning, which can cause harm to people. In addition, reuse or chaining of input signals (Sculley et al., 2015), poor patch dispatch (Sun et al., 2020), frequent software updates (EC, 2020a), and operating system faults (Hussain & Hong, 2020) are mentioned as

risk factors. A model developed for a CPS cannot be simply used for another CPS (McKee et al., 2017).

4.1.6.2 ANALYTICAL MODEL management risk factors

Describing reality is a major challenge in CPS design, particularly if it has autonomic functions and must respond quickly to safety-critical situations (McKee et al., 2017). There are several reasons a model can differ significantly from reality to make correct predictions (Pereira & Thomas, 2020). Reality can be complex, dynamic, unpredictable (Cabour et al., 2021; Coscia et al., 2021; Google, 2019) and not fully known (Pereira & Thomas, 2020). The model description of reality may originate from different information sources or have been developed for a smaller set of situations than those actually occurring in practice (Cabour et al., 2021). A model can differ from reality because nonlinear parameters are simplified (Ajayi et al., 2018), or because events or scenarios are omitted (Vogt, 2021; Bolbot et al., 2019; Dey & Lee, 2021). This can lead to failing system diagnosis and prognosis (Bolbot et al., 2019) or to persistent, yet undiscovered, mistakes the model does not flag up (Google, 2019). Poor model predictions lead to poor safety (Fosch-Villaronga & Mahler, 2021; Hussain & Hong, 2020).

4.1.6.3 MACHINE LEARNING management risk factors

Machine learning can fail in several ways. Machine learning faults can originate from several factors. The objective function may be incorrectly specified, inappropriate, incomplete or allowing 'objective drift' (Pereira & Thomas, 2020; Amodei et al., 2016; Google, 2019). A lack of clear requirements (EC, 2020a), limitations to avoid undesirable manners to achieve a goal (Dey & Lee, 2021) or a lack of realism in the objective (Google, 2019) may lead to unethical choices.

The ML algorithm drift concerns parameters that are not limited or safeguarded, for example, by not allowing a self-driving car to ignore safety warnings (Shneiderman, 2020), or by not stopping drift towards unacceptable action (Dey & Lee, 2021; Fosch-Villaronga & Mahler, 2021). ML algorithms can deliver incorrect outputs if fed by incomplete training data (Google, 2019), unbalanced or unrealistic real-world data (Pereira & Thomas, 2020), flawed sensor data, such as poor object recognition (EC 2020a), or a shifting situation over time, for example, a changing disease pattern or a change in the CPS setting (Challen et al., 2019; Leslie, 2019; Magrabi et al., 2019). If an 'outlier case' or a new 'edge case' is presented to the model, i.e. the parameter values extending outside the training data value range, the 'extrapolated' outcome may deviate significantly from reality (Magrabi et al., 2019). Some CPS application areas have many 'special cases' that need to be present in the training data, for example, for self-driving cars (Shneiderman, 2020).

The environmental conditions around a CPS may influence the CPS performance, for example, via temperature, fog, smoke, rain, poor lighting, awkward postures, electromagnetic interference, color pattern, object appearance, obstacles in the field of view, noise, or aggressive substances (Sculley et al., 2015; Soldatos & Kyriazis, 2021; Hussain & Hong, 2020; Khalid et al., 2017; Pereira & Thomas, 2020; Van Acker, 2020; EC, 2020b; Calvetti et al., 2020; Liu, 2022). Sculley et al. (2015) identified several ML risks that may be hidden in the system, software, and data. ML is currently the subject of research, and the question is how to handle ML in a safe and secure manner, for example, on the assurance of its functionality (Pereira & Thomas, 2020), dealing with accidents in the CPS (Amodei

et al., 2016), and new edge cases (Hamon et al., 2020). More understanding is needed regarding what else can be done to prevent accidents with machine learning systems (Amodei et al., 2016).

4.1.6.4 AI-APPLICATIONS management risk factors

AI applications belong to a separate category of software. AI application software can either be running on a local server in the direct vicinity of a CPS workplace or take part in the CPS as an online service delivered by an external contractor. This renders AI applications vulnerable to security breaches. Furthermore, AI applications can have complex interactions with other software applications in a CPS, may have a slow and therefore no longer real-time response from a remote source, can be mistaken, and can be temporarily unavailable, for example, due to an update or a network connectivity problem (Soldatos & Kyriazis, 2021).

4.1.7 DATA

4.1.7.1 QUALITY, MANAGEMENT, TRANSFER, PROCESSING & STORAGE management risk factors

Data quality may be affected by the high volume (McKee et al., 2017), poor coordination between data sources (Dey & Lee, 2021), diversity and non-structuredness (Angelopoulos et al., 2019), and differences between training, test, and final real-world datasets in the data validation process (Amodei et al., 2016; EC, 2020a).

More factors are relevant, such as faulty data (EC, 2020b), edge data, corrupted data, unforeseen cases (Dey & Lee, 2021), and poor, incomplete, inaccurate, unreliable, biased, or incorrect information (Leong, 2018; Soldatos & Kyriazis, 2021; Pereira & Thomas, 2020; Ajayi et al., 2018; Challen et al., 2019; Magrabi et al., 2019; Tschider, 2018; Huang et al., 2019; Leiber, 2017; Burggraaf et al., 2019). In the CPS development phase, engineers must prevent poorly chosen system training cases, the absence of rare cases, underrepresented cases, mixing of practice and theory data, bias in part of the data, too small training set data, non-realistic data, poor coverage by a training set of cases occurring in reality (Pereira & Thomas, 2020; Amodei et al., 2016; EC, 2020a; Challen et al., 2019), and extreme edge cases or input outlier cases (Leslie, 2019). These are all mentioned in the literature as quality problems that lead to unsafety.

Communication and data transfer quality may be affected by packet loss (Soldatos & Kyriazis, 2021; Peserico et al., 2021; Liu, 2022), delays or loss of information between system components (Bolbot et al., 2019; Olowononi et al., 2020), bit error rates and signal distortions due to multipath propagation, or by poor signal-to-noise ratio in sensor data (Liu, 2022; Khalid et al., 2018; Peserico et al., 2021). In practice, the accuracy of data transfer is no higher than 98% (Liu, 2022). Transmission is important, and processing and storage also affect results (Liu, 2022).

4.1.7.2 PROTECTION management risk factors

Not only the data itself but also their management is important; for example, unavailability of the origin, context, or background of data (Magrabi et al., 2019; Wang & Wang, 2021), or being unprotected (Hamon et al., 2020). Data must be protected, for example, to avoid theft, to discover and identify new risks, to safeguard the rights and interests of users, for example, privacy, and to keep information up to date (Hamon et al., 2020; Tschider, 2018; Wang & Wang, 2021)

4.1.8 OPERATIONS

4.1.8.1 MAKING A PRODUCT, EXECUTING A TASK management risk factors

Over time, a wide variety of CPS applications have been developed, for example, in the manufacturing industry, healthcare, smart buildings, transportation, and other areas. (Jiang et al., 2019; Mezzanotte, 2019).

During operations, CPS functions are influenced in real time by human worker and operator behavior, site environmental conditions, resource availability, and the performance of the other system components.

Human workers in a fully automated CPS production environment stand out as a risk because of their high failure rate owing to human error, variable appearance that challenges CPS sensors and recognition software, sometimes incomprehensible voice communication, and unexpected movements or gestures (Bolbot et al., 2019; Van Acker, 2020; Soldatos & Kyriazis, 2021; Fosch-Villaronga & Mahler, 2021; Khalid et al., 2017; Hussain & Hong, 2020).

Changes in environmental conditions can influence CPS performance, such as temperature, noise, weather conditions, and poor electromagnetic cleanliness (EMC).

Both external resources and system components can fail at any time. This is of paramount importance, both for product or service quality and for the physical safety of workers and patients (Costantino et al., 2021; Khalid et al., 2017).

4.1.9 MAINTENANCE

4.1.9.1 DEPENDABILITY management risk factors

The dependability of a CPS is a safety concern because of the human presence and impact or crush risks.

Disturbances and failures, both during normal operations and during testing, can originate from accuracy and repeatability of the end effector position (Haidegger et al., 2020), the reliability and variability e.g. in communication protocols (Hussain & Hong, 2020; Hamon et al., 2020; Furstenau et al., 2022), unreliable or inaccurate design (Liu, 2022; Costantino et al., 2021), new or customized system components not performing as expected (Liu, 2022), uncertainty (Renaud et al., 2021), incomplete testing, verification and evaluation of the CPS design during development (Pereira & Thomas, 2020; Amodei et al., 2016; Google, 2019; Magrabi et al., 2019; Bolbot et al., 2019), poor robustness e.g. while subjected to differences between training data and real-world data, also referred to as ‘distributional shift’ (Dey & Lee, 2021, Amodei et al., 2016), from poor design choices in fault tolerance and fail safe characteristics (Pereira & Thomas, 2020; Amodei et al., 2016; Leslie, 2019; Koene et al., 2018), interrupted connectivity (Begic & Galic, 2021), wearable sensors (Calvetti et al., 2020) and from poor preparedness for cyber-attacks (Olowononi et al., 2020).

4.2 CPS Safety and security management system (SSMS) aspects

An SSMS is primarily designed to prevent unwanted events and harm to people and the environment by controlling the risks that exist in an organization. These systems must deal with CPS risk factors.

Because there are many alternative SSMS designs, and the security and safety aspects may or may not be integrated in a single system, the authors adhere to the basic functions of such systems (Li & Guldenmund, 2018). Following the safety management system structure used in the European high-risk chemical industry and in hospitals (EC, 2012) is considered practical here. These SSMS systems consist of several sections, as listed in Table 2.

Section	Aspects covered
policy	Safety and security policy matching the hazards, activities and complexity
1-people	Organization and personnel, training of all people involved
2-risks	Identification of hazards and appraisal of risks
3-safe work	Safe work during operations and maintenance for all people involved
4-change	Change management taking safety into account
5-emergency	Emergency preparedness, training for emergency situations
6-metrics	Monitoring safety system performance, (near) accident investigations
7-review	Regular review and adjustment of safety system performance

Table 2 - SSMS sections and the aspects they cover

Several CPS-related challenges on safety and security systems in industry and healthcare are identified from the literature findings and are allocated to the SSMS sections below.

4.2.1 SSMS section Safety and Security policy

4.2.1.1 Attitudes of stakeholders

The attitude of people involved, varying from approaching Big Data and IoT as only technical, the engineers' focus on the question 'does it work' (Costantino et al., 2021; Mezanotte, 2019; Spiekermann & Winkler, 2020), negativity towards the cyber security aspect (Renaud et al., 2021; EC, 2020a), regarding humans as a problem (Stefana & Paltrinieri, 2021; Spiekermann & Winkler, 2020), operators mistrusting the system (Fletcher & Webb, 2017; Coscia et al., 2021), overreliance on technology (Bolbot et al., 2019; Costantino et al., 2021; Shneiderman, 2020), ignoring CPS ethics aspects beyond mere machine ethics (Trentesaux & Karnouskos, 2019), ignoring the safety aspect altogether, e.g. when developing or purchasing a surgical robot system (Liu, 2022), routine users complacency or deterioration of worker skills (Google, 2019; Fletcher & Webb, 2017; Laurent & Fabiano, 2022), lack of business development support (EC, 2020a; Mohammadpoor & Torabi, 2020), a defensive attitude of personnel not provided with an informed choice (Fletcher & Webb, 2017) or not kept informed of the changes for them when a CPS is installed (Costantino et al., 2021). All these attitude problems can impede the development of safer and more secure CPS.

4.2.1.2 Safety and Security become entangled

A CPS requires a clear focus on occupational safety (Deschacht, 2021; Mezzanotte, 2019; Vinuesa et al., 2018; Laurent & Fabiano, 2022), workers' physical and mental condition (Costantino et al., 2021), risk awareness (Cantelmi et al., 2021), new risks originating from new technology (Lindhout, Kingston & Reniers, 2019; Costantino et al., 2021) and guarantees a sufficiently high level of safety

(Peserico et al., 2021; Olszewska et al., 2020; McKee et al., 2017). Pereira and Thomas (2020) point out that setting such an acceptance level leaves open the possibility of an accident if its risk is just below the acceptance threshold level.

A CPS also requires a clear focus on cyber security (Mezzanotte, 2019; Ogbuke et al., 2022; Marjumin et al., 2019; Malhotra et al., 2021; Tschider, 2018; Vinuesa et al., 2018; Furstenau et al., 2022) in order to avoid harm to people and damage (Renaud et al 2021; Peserico et al., 2021; EC, 2020b; Amodei et al., 2016; Liu, 2022; Hamon et al., 2020) by big and strong moving machines (Soldatos & Kyriazis, 2021). This is needed specifically for software development (Reniers & Amyotte, 2012), interaction between multiple AI systems (Soldatos & Kyriazis, 2021, EC, 2020b), securing data integrity (Soldatos & Kyriazis, 2021; Fosch-Villaronga & Mahler, 2021; Khalid et al., 2018; Olowononi et al., 2020; Dey & Lee, 2021; Google, 2019; Liu, 2022), for cyber-attacks by e.g. hackers, virus or malware (Peserico et al., 2021; Bellini et al., 2021; Angelopoulos et al., 2019; Soldatos & Kyriazis, 2021; Bolbot et al., 2019; Fosch-Villaronga & Mahler, 2021; Hussain & Hong, 2020; Khalid et al., 2018; McKee et al., 2017; Olowononi et al., 2020; Jing et al., 2014) and for abuse such as entering incorrect data affecting the system performance (Dey & Lee, 2021; Van Acker, 2020; Amodei et al., 2016; Sculley et al., 2015; Google, 2019; Leslie, 2019).

Consequently, a cybersecurity breach may lead to physical injury. The relationship between safety and security is changing (Reniers & Khakzad, 2017). Traditionally, safety and security have been separate disciplines in organizations. Because cyber risks affect everyday operations in a CPS setting, safety and security disciplines are interconnected (Chehri et al., 2021; Hopcraft et al., 2021; Matta et al., 2021; Huang et al., 2019; Fosch-Villaronga & Mahler, 2021; Begic & Galic, 2021; Reniers & Amyotte, 2012).

4.2.1.3 Risk acceptance criteria

A CPS as installed and the procedures to operate it safely need to deliver sufficiently high safety levels (Peserico et al., 2021). This makes several criteria important for safety and security policy statements and their implementation in safety management systems. Such criteria are e.g. safe operating window parameter values (Pereira &Thomas, 2020), an acceptable collision impact injury risk level (Fletcher & Webb, 2017), a maximum force level exerted on a person when hit by a robot taking into account differences between body parts (Khalid et al., 2017), components that satisfy SIL-3 level requirements (Peserico et al., 2021), the definition of a ‘safe state’ or ‘base line’ risk level that may not be exceeded (Pereira & Thomas, 2020; Fosch-Villaronga & Mahler, 2021) and what level of evidence is acceptable in support of a decision (Magrabi et al., 2019).

4.2.1.4 Decision making

The idea of using both real-time data and big data analytics for safety and security management purposes has existed for quite some time (Mol, 2003; Kirwan, 2008; Uraikul et al., 2007). However, taking safety and security decisions automatically is currently considered a contentious, complex, and risky point, and there are only a few practical applications. The authors contend that further research is necessary.

4.2.2 SSMS section 1-People

4.2.2.1 Employee knowledge and skills

BD, AI, and IoT components make a CPS difficult to understand, not for those designing it, not for those operating it, and not for those in charge. Keeping CPS and the people around it safe implies the need for specialized knowledge and skills. In the current labor market, such expertise is difficult to obtain, while commercial forces bring about an increasing demand for trained, specialized personnel. Training is a challenge because explainability, understanding of the complex CPS work environment, newness for training specialists, dynamic character, and frequent changes are the main concerns (Patriarca et al., 2021; Leiber, 2017; Leong, 2018; Fletcher & Webb, 2017; Bolbot et al., 2019; Coscia et al., 2021; Mohammadpoor & Torabi, 2020).

Qualified employees are a key requirement for the safe operation of CPS, but it is increasingly difficult for organizations to find and train these people. The increasing need for new employee competencies (Costantino et al., 2021), change in employee functions, growing need for data scientists and data engineers (Laurent & Fabiano, 2022), absence of systematic knowledge sharing (Johnsen et al., 2009), and lack of attention to adapt to new knowledge (Magrabi et al., 2019) are also concerns. Big Data and IoT knowledge and skills are needed for workers, operators, and safety engineers, and “data literacy” is being suggested as a prerequisite for human resource management (HRM) prerequisite (Wang & Wang, 2021, p5). Attention must be paid to the mismatch between existing operator knowledge and experience and the newly required skill set (Laurent & Fabiano, 2022). The knowledge and skills of safety and security engineers need to be extended via training beyond the currently existing health and safety domain boundaries.

Training needs to explicitly address basic knowledge of BD, AI and IoT components in the system (Begic & Galic, 2021; Bolbot et al., 2019), multidisciplinary team work (Bolbot et al., 2019), unwanted event scenario’s (Zimmerman & Renaud, 2019; Johnsen et al., 2009), emergency preparedness (Angelopoulos et al., 2019; Johnsen et al., 2009), safe behavior in the presence of robots (Fletcher & Webb, 2017; Costantino et al., 2021; Liu, 2022), scenario training (Johnsen et al., 2009), specific safety management (Laurent & Fabiano, 2022) and ‘data literacy’ (Wang & Wang, 2021).

The fact that ML data-led models are of a different kind than law-based models, and hence, it is more difficult to discuss data-led models in a meaningful way via reasoning, formulas, and parameters (Magrabi et al., 2019; Leiber, 2017), makes it important for safety engineers to be aware of this.

4.2.2.2 Understanding and awareness

Not only is the human understanding of CPS as a system important. Additionally, the reverse is important.

CPS must also be sufficiently enabled to sense human whereabouts, behavior, movements, sounds, and emotions (Sun et al., 2020; Fosch-Villaronga & Mahler, 2021). Poor system understanding can originate from unexpected human appearance (Medina et al., 2019). Poor communication between the CPS and operator may lead to insufficient situational awareness (Bolbot et al., 2019) for both robots and humans.

Awareness goes hand-in-hand with the safety and security of CPS. An update in part of the software, for example, adding or improving functionality or new data, must be regarded as a safety-critical activity because not all of the possible algorithm or machine learning outcomes can be foreseen (Bolbot et al., 2019; Amodei et al., 2016; Hamon et al., 2020). This requires special attention during training (Fletcher & Webb, 2017).

Gathering real time safety data and using this real-time in scientific predictive analytical models, also referred to as "computational safety," allows safety management to preview e.g. hazardous situations, and provide early warnings (Mol, 2003; Uraikul et al., 2007; Leveson & Stephanopoulos, 2014; Blokland & Reniers, 2019; Wang & Wang, 2021; Wang, 2021).

4.2.3 SSMS section 2-Risks

4.2.3.1 Risk assessment

The design of a potentially hazardous system should be based on sound risk assessment (ISO-31000; Spiekermann & Winkler, 2020). New hazards emerge from a range of new cyber, technical, and human system properties and from the way such systems are designed, developed, operated, and maintained. Reasonably, foreseeable use and misuse by employees and unauthorized intruder activities must be included in the safety risk inventory (EC 2020b). In itself, the large number of potentially dangerous combinations of conditions or events during normal operations (Ale et al., 2014) poses a challenge for CPS risk assessment. As Trentesaux and Karnouskos (2019, p9) put it: "calculating all consequences of an action is impossible." In addition, concerns have been raised regarding whether safety management as a professional discipline can keep up with the pace of emerging AI, BD, and IoT technologies, including all of its new associated risks (Pillay, 2015). The new element of cooperating robots and humans in the same workplace further extends the risk inventory even further (Jafari et al., 2022; Zheng et al., 2017; Johnsen et al., 2009; Medina et al., 2019; EC, 2020a).

4.2.3.2 Risk identification

Identification of such new risks, let alone a way of controlling these new risks, cannot be done based only on previous experience and learning from accidents, nor can they be predicted from extrapolation (Pawson et al., 2011). Risk identification requires new thinking, needs imagination, and daring to investigate the unknown, organizations need 'risk appetite' (Gjerdrum & Peter, 2011). New types of occupational risk scenarios on a joint robot-human shop floor come from different angles. A robot may simply not be able to see a human crossing its path owing to sensor failure (Haidegger et al., 2020). This – and many other possible scenario's – must be included in CPS software, e.g. as a possible 'event,' and be linked to evasive, corrective and emergency CPS actions, also included in the software. An example of an 'event' not taken up in the 'event space,' was 'a person falling down' in front of an automated vehicle, which was leading to a lethal accident (Vogt, 2021). Poor ergonomics, for example, in the user/operator-system interface, can lead to unintended or unexpected robot movements, causing harm to people, such as collisions (Moencks et al., 2022; Medina et al., 2019; Khalid et al., 2017; Olowononi et al., 2020).

Occupational safety, specifically workplace physical safety, can be affected not only by CPS design but also by CPS failure (Medina et al., 2019; Costantino et al., 2021; Tschider, 2018; Dey & Lee,

2021). Mental factors, such as stress, excessive workload, and physical or mental fatigue, need to be considered as well (Costantino et al., 2021; Van Acker, 2020).

4.2.3.3 Health care

In healthcare, patient safety needs special attention because wearing a device, skin damage, machine learning drift, or application on an unanticipated patient may lead to harm to vulnerable people (Fosch-Villaronga & Mahler, 2021; Liu, 2022; EC, 2020b). Robots interacting with vulnerable people must not cause feelings of losing control of one's life (Challen et al., 2019) or being treated like an object (Olszewska et al., 2020). This also applies to not following operating instructions, misunderstandings within a medical team, and errors in handling CPS medical devices (Liu, 2022). In the chemical industry one might ask whether detection of unsafety can help to control new and possibly varying risks (Leveson & Stephanopoulos, 2014), e.g. via more artificial intelligence and sensor technology continuously generating 'safe state' warnings (Uraikul et al., 2007), a "big-data nervous system" approach (Lindhout et al., 2020, p4), pro-active instantaneous measuring of (un)safety conditions (Blokland & Reniers, 2019) or via dynamic risk assessments (Mol, 2003). Current risk identification methods are not designed for the latest CPS developments. Multi-method risk identification is required to assess hazards, risks, and scenarios (Bolbot et al., 2019; Pereira & Thomas, 2020). Risk assessment must consider intended use, foreseeable use, and, where applicable, reasonably foreseeable misuse (EC, 2020b).

4.2.3.4 Safety culture

Hopcraft et al. (2021) wondered how the 'cyber' and 'human' aspects can both be integrated into safety culture. The "The Egg Analythical Model" (TEAM) integrated safety culture model (Vierendeels, 2018) already includes the 'human' aspect and can be used as a starting point for the introduction of the 'cyber' aspect. The observed integration of safety and security' aspects in the operations of a CPS found in this study also constitutes an argument for the introduction of the 'security' aspect in the TEAM model. The CPS aspects of shared responsibility, awareness (Renaud et al., 2021), trust (Angelopoulos et al., 2019) and protection from physical and mental harm (Fletcher & Webb, 2017) match the existing TEAM model content.

4.2.4 SSMS section 3-Safe work

4.2.4.1 Workplace safety

For workers, operators, engineers, and managers, the key factors for safe and secure operations are their understanding of the CPS as a system (Soldatos & Kyriazis, 2021; Patriarca et al., 2021), facilitated by ergonomic human – machine interfaces (Marjumin et al., 2019; Costantino et al., 2021; Liu, 2022), their awareness of the intricacies and risks of both human and robot presence on the shop floor (Patriarca et al., 2021) and the need for clear two-way communication between humans and robots (Angelopoulos et al., 2019).

Factors contributing to safety include working in open, joint human and robot, collaborative working spaces (Jafari et al., 2022; Trentesaux & Karnouskos, 2019), working without necessary personal protection (Costantino et al., 2021; Van Acker, 2020), working under pressure (Van Acker, 2020), communication errors (Hussain & Hong, 2020), changes such as new procedures, roles, or digital

tools (Costantino et al., 2021), and a high cognitive load (Van Acker, 2020). Workers feel that their daily interpersonal contact is threatened (Olszewska et al., 2020).

4.2.4.2 Human error

Any physical presence of humans in the direct vicinity of a working CPS has thus far often been avoided by fencing them out. This has been the primary safety strategy for several decades. Recent developments have led to communication between cobots and co-workers, back and forth, physical interaction, and cooperation in production environments (Bolbot et al., 2019). Previously, human error could be hidden in data entry (Pereira & Thomas, 2019), in the model of reality, or in software; in recent CPS designs, human error can also be present in the direct interaction with a cobot (Bolbot et al., 2019; Soldatos & Kyriazis, 2021). Human error can have severe safety consequences for workers. Human error can occur due to limited reliability and known error rates of a human in the interaction (Bolbot et al., 2019; Van Acker, 2020), lack of knowledge or skills, poor coordination between operators (Medina et al., 2019), and poor workplace situations. More insidious, and certainly less intuitive, is the vulnerability to loose clothing, new colors, patterns, or long hairstyles. Also a CPS, equipped with its array of sensors, and programmed how to 'look' at the people around it, can get confused and let a cobot make unexpected movements (Medina et al., 2019). Similarly, incorrectly interpreted movements or body positions, unintelligible voice commands, and unrecognized gestures (Fosch-Villaronga & Mahler, 2021; Khalid et al., 2017) may cause unsafe conditions, such as collisions, impacts, or entrapment risks. After a cobot suddenly hurts or grabs a worker, the risk for other workers coming to the rescue that they too can get hurt is clear, a.o. from various video's about 'industrial robot accidents' circulating on the internet. This underlines the necessity of training CPS in accident scenarios.

4.2.5 SSMS section 4-Change

4.2.5.1 Management of change

Changes, for example, in the workplace, operating procedures, ergonomic interfaces, applied technology (Costantino et al., 2021; Deschacht, 2021), human co-worker needs (Yilma et al., 2021), or by combining new with existing systems (McKee et al., 2017), all require change management. The changes and their potential consequences vary considerably. Change management is necessary between the CPS training and operational phases. For example, sensor replacement, taking place before the start-up of CPS operations, can invalidate the system training data invalid (Pereira & Thomas, 2020). After start-up, the slow" lt' as-built system evolution over time via software updates, maintenance and repairs, product changes, and sensor improvements may have similar effects (Bolbot et al., 2019; EC, 2020a; EC, 2020b). Such effects can also be caused by trends in disease patterns, patient group characteristics, and treatment options (Magrabi et al., 2019; Challen et al., 2019).

4.2.5.2 Use on other location

A CPS is designed for a task at a specific location. This is important for both the dataset used as a basis for operations and for the algorithms used. This is especially true for machine learning (ML) applications because they may involve a combination of high-precision geometrical settings, alignment, lighting conditions, sensor settings, and other factors that depend on location-tied peculiarities. Hence, moving a CPS to another location can make the ML model dataset no longer

applicable (Pereira & Thomas, 2020) because the new reality differs from the old and trained reality. Unsuspected actions, such as raising the light intensity level – revealing what was hidden in a shadow or creating a salient reflection – can place (part of) the operational real-world data outside the boundaries of the ML model training dataset (Van Acker, 2020). Thus, the ML model will not be able to usefully respond to the 'new' part of the data. A change, such as the inclusion of an aspect that was not present in the initial CPS risk assessment or a change in the specifications of the product that is processed by the CPS, can also have safety effects (EC 2020a; EC 2020b). Change may also compromise the safety of human workers because it can affect the collaboration quality between cobots and co-workers (Yilma et al., 2021). Changes may originate from the outside world in contact with CPS via, for example, online services and affect safe operation (Sculley et al., 2015). The involvement of workers in the change process is not only essential for their safety and awareness, but also to avoid job insecurity (Costantino et al., 2021).

4.2.6 SSMS section 5-Emergency

4.2.6.1 Emergency preparedness

Johnsen et al. (2009) and Angelopoulos et al. (2019) identified poor scenario training as a main shortcoming of emergency preparedness in CPS settings. Shneiderman (2020) and Bolbot et al. (2019) pointed out the importance of the thorough identification of hazards, risks, and accidents based on emergency scenarios. Moreover, they underline the importance of paying sufficient attention to a plethora of possible special cases.

4.2.6.2 Automatic versus human action

Automatic safety devices have millisecond response times, precluding any human intervention to stop the automatic action (Peserico et al., 2021, p.11,12), even if it is based on incorrect or incomplete data. Preventing harm to people requires real-time system response (Olowononi et al., 2020). However, real-time data gathering from sensors via a bus system, data exchange between interacting systems, and online cloud services are as good as the slowest communication cycle polling time, sensor response time, encryption key, network overhead functionality, or RFID system interaction. In practice, this value is 125 µs or more (Huang et al., 2019; Peserico et al., 2021; Tramarin et al., 2019; Soldatos & Kyriazis 2021; Khalid et al., 2017; Jing et al., 2014; Liu, 2022).

False alarms caused by delays in data exchange or faulty sensor outputs can trigger automatic but potentially unsafe actions (Liu, 2022; Dey & Lee, 2021). In this way, an autonomous response of an AI system may cause harm to people owing to incorrect or unintended calculation outcomes (EC, 2020b).

4.2.7 SSMS section 6-Metrics

4.2.7.1 Safety metrics

As in any industrial or healthcare organization, 'safety' must be measured in a CPS setting. Real-time information regarding worker safety is needed. This is compared to the chemical process industry, where process parameter measurement provides real-time information to operators in the control room, and where the production process currently is with respect to the safe operating window. To this end, safety metrics, especially those for relative speed and distance between humans and robots,

need to be reliably and continuously calculated in real time in areas where both the cobot and co-worker share the same workspace (Khalid et al., 2017; Fletcher & Webb, 2017).

Gathering and analyzing safety data in support of decision making by safety management about actions to be taken (Kirwan, 2008; Wang et al., 2021), is referred to as "safety intelligence". This could be an automatic action or monitoring of workers or patients (Wang & Wang, 2021).

Because cobot-to-co-worker distance safety metrics have become life-savers, the authors contend that they are safety critical. Hence, they must be treated as a different category and separated from Safety Performance Indicators (SPI) and other traditional safety performance metrics (Swuste et al., 2016).

In practice, distance metrics are not always present, verified, or tested however (Pereira & Thomas, 2020; Khalid et al., 2017). Although proven designs for CPS safety metrics are available (Haidegger et al., 2020, p267-268; Koene et al., 2018), the implemented safety and security metrics are not always adequate for the measurement of human performance (Pereira & Thomas, 2020) or can be vulnerable to incorrect interpretations (Pasman & Rogers, 2014). Safety critical metrics, built from sensors, connections, and software algorithms, can not only be compared to key performance indicators (KPI) but also to process safety performance indicators (SPI), and should therefore meet special requirements (Khan et al., 2009; Swuste et al., 2016).

The time required by a system to detect an obstacle must be maintained at the criterion level to ensure safety (Pereira & Thomas, 2020).

4.2.7.2 Security aspects

In a CPS, there is an interaction with security because a cyber-alarm may overrule all regular CPS operations, including real-time metrics for the protection of people (Hopcraft et al., 2021). Taking up such security aspects in the software can bring back control over alarming situations (Reniers & Amyotte 2012). A special concern is machine learning disturbance caused by irregular interruptions, such as accidents or deliberate misuse incidents. Not only the changed safety situation (Google, 2019) but also both machine learning and learning from accidents require attention (Amodei et al., 2016; Ajayi et al., 2018).

4.2.7.3 Safety information

Gathering legal, standard, accident reporting, and regulatory data, usually referred to as "safety information," allows safety management to comply with the requirements (Wang & Wang, 2021). Applications could address e.g. learning from accident data via sharing (Zheng et al., 2017; Esreda, 2015), risk inventory (Pereira &Thomas, 2020; Bolbot et al., 2019; Lindhout, Kingston & Reniers, 2019; Lindhout et al., 2020) and real time visualization techniques, e.g. "safety dash-board" (Wang & Wang, 2021).

4.2.8 SSMS section 7-Review

Very few sources mention the need for an evaluation or management review of safety and security, as achieved in settings with CPS and humans on the same shop floor. Such an evaluation and review process is important in the operational phase of CPS. Ideally, it stretches out over all life cycle phases, including design, development, testing, operational use, and ongoing surveillance (Magrabi et al.,

2019). The authors contend that regular management reviews of safety and security, and taking action are as important for a CPS as they are for any other safety-critical activity.

4.3 Government and regulatory activities

Accou and Reniers (2019) wonder whether government regulator, safety, and insurance inspectors can sufficiently ascertain what goes on in a CPS. One of the obstacles is that " . . . transparency . . . can negatively interact with profit and speed of production" (Spiekermann & Winkler, 2020, p1).

Associated problems here are both overdue and outdated legislation and the absence or non-binding status of standards (Costantino et al., 2021) in the presence of a need for law enforcement (Renaud et al., 2021) in view of the already existing use of robots in settings with humans capable of inflicting harm to people, such as in personal care (Fosch-Villaronga & Mahler, 2021).

There are several additional issues, for example, the limited possibilities for regulators to intervene because of a lacking legal framework, settings where a CPS is not confined to a single site or country; opacity, that is, lacking AI or ML systems transparency and poor explainability of what goes on in a CPS (Soldatos & Kyriazis, 2021; Dey & Lee, 2021; Coscia et al., 2021; EC, 2020a; EC, 2020b; Tschider, 2018; Vinuesa et al., 2018; Leiber, 2017; Koene et al., 2018; Spiekermann & Winkler, 2020), the presence of black boxes (EC, 2020b; Magrabi et al., 2019; Koene et al., 2018), a lack of regulator inspector knowledge; and a lack of technical means to carry out an inspection (Tschider, 2018; EC, 2020a).

5 Discussion

5.1 Finding issues

Each admitted literature source was screened for text parts relevant to the sub-questions. The relevance of a text part is determined by mentioning a problem, concern, or issue (Timulak, 2014) in connection with safety or security and CPS technology and operations. With this qualitative method, a rough first indication can be obtained regarding the relative importance of the many issues found in the literature in this study. To this end, we gathered 864 text extracts related to safety and security, each mentioning a particular issue. These extracts were then compared to identify similarities and logical combinations between the issues. Next, using simple text analysis and meta-synthesis (Noah, 2017), the data were condensed into 53 clusters. These clusters are ranked in Table 3 based on the frequency mentioned.

The first observation here is that Table 3 shows that the five most frequently mentioned issues are: "risks"(56), "data quality"(49), "workplace"(44), "security focus"(39), and "software quality"(35). These clearly highlight key concerns.

The second observation is that there are several other issues. Although each mentioned less frequently, they clearly outweigh the five most frequently mentioned ones.

Issue Nr	Description	Frequency count	Group code	Issue Nr	Description	Frequency count	Group code	Issue Nr	Description	Frequency count	Group code
1	risks	56	2	21	transparency	18	R	41	verification	8	F
2	data quality	49	V	22	hardware	17	V	42	discrimination	7	E
3	workplace	44	3	23	safety focus	17	P	43	dynamic	7	I
4	security focus	39	P	24	attitude	15	P	44	patient	7	V
5	software quality	35	V	25	understanding	15	1	45	objective	6	V
6	human factor	28	V	26	describing reality	14	V	46	only technical	6	P
7	comm-transfer	27	V	27	metrics	14	6	47	user i/f	6	3
8	cyber security	26	V	28	terminology	14	1	48	emergency	5	5
9	sensors	24	V	29	regulator	13	R	49	maintenance	5	V
10	design	22	D	30	diagnostics	11	V	50	automatic	4	5
11	organization	22	P	31	reliability	11	F	51	accuracy	3	F
12	privacy	22	E	32	trend	11	P	52	uncertainty	3	F
13	change	21	4	33	abuse	10	V	53	evaluation	2	7
14	ethics	21	E	34	criteria	10	P				
15	standards	20	S	35	governance	10	G				
16	training	20	1	36	resources	10	V				
17	decision making	19	7	37	safety culture	10	2				
18	machine learning	19	V	38	big data	9	I				
19	complexity	18	D	39	accident handling	8	6				
20	environmental cond.	18	V	40	human error	8	3				

Table 3 – 53 CPS safety and security issues, frequency counts and cluster group codes

5.2 Key areas of concern

The 53 clusters were screened based on whether they relate to either the CPS itself or to safety and security management. To obtain a better 'grounded' overview of where the main problem areas are, a second analysis is performed. To this end, 53 safety and security aspect clusters and their frequency counts are allocated to eight CPS aspect groups (see figure 1) and eight SSMS section groups (see Table 2), using the descriptions in sections 4.1 and 4.2, respectively. The allocation of the 53 issues to these 16 groups is indicated by the group codes in Table 3. Frequencies for each issue were added to calculate the respective group frequencies. Using this, an indicative frequency histogram was generated, as shown in Fig. 2.

The results of this analysis demonstrate how the findings relate to safety and security management.

The first observation is that the attention to new or increased vulnerabilities stands out, with 306 out of 864.

The second observation is that, in the current state of CPS development, attention is most frequently focused on the CPS system and its constituent parts. This is shown by 498 counts out of a total of 864.

The third observation was that the remaining 366 counts were allocated to SSMS aspects. The main part, 120 of these 366, is related to safety and security policies.

The fourth observation is that, out of the seven SSMS sections, the first three have higher counts than the last four sections. The Risks are most frequently mentioned when it comes to safety management

systems, followed by personnel, training, and safe work. Change management, emergency preparedness, safety and security Metrics and Management reviews are mentioned less frequently.

Group code		Frequency counts
	CPS aspects	
V	new or increased vulnerabilities	306
E	ethics	50
D	systems complexity & design	40
R	government and regulatory activities	31
F	system dependability	25
S	standards	20
I	interactivity, dynamics, connectedness	16
G	governance	10
		498
	Safety and security management system sections	
P	Safety and security policy	120
1	People	49
2	Risks	66
3	Safe work	58
4	Change	21
5	Emergency	9
6	Metrics	22
7	Review	21
		366
Total counts		864

Figure 2 - Frequencies of safety and security issue clusters per aspect group code

5.3 Limitations to this study

The findings in this study are robust because of the number of recent scientific, secondary, and tertiary sources underpinning the results (Table 1) and the straightforward analysis (Section 5.) that uses text extracts and a meta-synthesis method to avoid bias.

The CPS-related hazards inventory presented in this study will be useful as a starting point for safety and security professionals when assessing CPS risks while building and running a Safety and Security Management System (SSMS).

With the development of Cyber Physical Systems (CPS), new design concepts have continuously emerged. Therefore, new and currently unknown hazards may occur over time. Therefore, the results of this study can be incomplete and must be regarded as indicative of current and future CPS concepts.

6 Conclusions and recommendations

Many challenges are faced by safety and security management in organizations using cyber-physical systems (CPS). This study explores these challenges and presents findings on both the general CPS aspects and aspects of safety management systems.

Due to the method chosen in this study and the high rate of change in the field explored, the findings presented must be considered indicative, probably incomplete, and subject to change.

Safety and security management must ensure the safety and security of workers, as robots have started to join them in common workspaces. Not only is support from standards and guidance lagging behind during the design of increasingly rapidly developing new CPS concepts. In addition, the external supervision of such new concepts implemented in an organization, as well as the regulatory role, struggles to keep up. So far, humans have received little attention in the development of complex high-tech manufacturing systems. A new type of safety-critical device is emerging: sensors watching human-robot interactions, human whereabouts, behavior, and emotions.

The authors consider the following recommendations to be of paramount importance for safety and security management in an organization working with a CPS:

-Implementation of an integrated SSMS in practice

-Development of guidance for the development of a CPS, considering safety, security, and ethics

-Further research on real-time applications in support of safety and security management

-Research and development of sensor arrays capable of observing human presence, behavior, and emotions.

-allocation of a special criticality status for metrics related to CPS-human interaction.

The practical application of these findings in safety and security management systems will contribute to avoiding harm to people and the environment in companies operating a CPS.

CRediT authorship contribution statement

Conceptualization: P.L. and G.R.; methodology: P.L.; writing—original draft preparation: P.L.; writing—review and editing: P.L. and G.R. All authors have read and agreed to the published version of the manuscript.

Funding or Grant

This research did not receive any specific grants from funding agencies in public, commercial, or

not-for-profit sectors.

Declaration of competing interests

None.

Use of AI

During the preparation of this work, the authors used no AI.

References

Accou, B., & Reniers, G. (2019). Developing a method to improve safety management systems based on accident investigations: the SAfety FRactal ANalysis. *Safety Science, 115*, 285–293. https://doi.org/10.1016/j.ssci.2019.02.016

Ajayi, A., Oyedele, L., Delgado, J. M. D., Akanbi, L., Bilal, M., Akinade, O., & Olawale, O. (2018). Big data platform to predict health and safety accidents. *World Journal of Science, Technology, and Sustainable Development, 16,* 2-21. https://doi.org/10.1108/WJSTSD-05-2018-0042

Akiyama, K., Alberdi, A., Alef, W. et al. (2022). First Sagittarius A* Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole in the Center of the Milky Way. *The Astrophysical Journal Letters, 930(2).* [The EHT Collaboration et al. et al 2022 ApJL 930 L12] https://doi.org/10.3847/2041-8213/ac6674

Ale, B., Van Gulijk, C., Hanea, A., Hanea, D., Hudson, P., Lin, P.H., & Sillem, S. (2014). Towards BBN based risk modelling of process plants. *Safety Science, 69*, 48–56. https://doi.org/10.1016/j.ssci.2013.12.007

Ali, N., Hussain, M., & Hong, J. E. (2020). Analyzing safety of collaborative cyber-physical systems considering variability. *IEEE Access, 8*, 162701-162713. https://doi.org/10.1109/ACCESS.2020.3021460

Amodei D., C. Olah, J. Steinhardt, P. Christiano, J. Schulman, & D. Mane (2016). Concrete problems in ai safety, arXiv preprint arXiv:1606.06565. https://doi.org/10.48550/arXiv.1606.06565

Angelopoulos, A., Michailidis, E. T., Nomikos, N., Trakadas, P., Hatziefremidis, A., Voliotis, S., & Zahariadis, T. (2019). Tackling faults in the industry 4.0 era—a survey of machine-learning solutions and key aspects. *Sensors, 20*(1), 109. https://doi.org/10.3390/s20010109

Arden W, Brillouët M., Cogez P., Graef M., Huizing B., & Mahnkopf R. (2010). "More-than-Moore". White Paper. IRC-ITRS roadmap. https://www.itrs2.org/uploads/4/9/7/7/49775221/irc-itrs-mtm-v2_3.pdf [Accessed April 4, 2022]

Asaad, R. R., & Abdulhakim, R. M. (2021). The Concept of Data Mining and Knowledge Extraction Techniques. *Qubahan Academic Journal, 1*(2), 17-20. https://doi.org/10.48161/qaj.v1n2a43

Assis-Dornelles, J. de, Ayala, N. F., & Frank, A. G. (2022). Smart Working in Industry 4.0: How digital technologies enhance manufacturing workers' activities. *Computers & Industrial Engineering, 163*, 107804. https://doi.org/10.1016/j.cie.2021.107804

Begić, H., & Galić, M. (2021). A Systematic Review of Construction 4.0 in the Context of the BIM 4.0 Premise. *Buildings, 11*(8), 337. https://doi.org/10.3390/buildings11080337

Bellini, E., Marrone, S., & Marulli, F. (2021). Cyber Resilience Meta-Modelling: The Railway Communication Case Study. *Electronics, 10*(5), 583. https://doi.org/10.3390/electronics10050583

Benna, M.K., & Fusi, S. (2016). Computational principles of synaptic memory consolidation. *Nature Neuroscience, 19*(2), 1697-1706. https://doi.org/10.1038/nn.4401

Bharadwaj, H. K., Agarwal, A., Chamola, V., Lakkaniga, N. R., Hassija, V., Guizani, M., & Sikdar, B. (2021). A review on the role of machine learning in enabling IoT based healthcare applications. *IEEE Access, 9*, 38859-38890. https://doi.org/10.1109/ACCESS.2021.3059858

Blokland, P., & Reniers, G. (2019). Measuring, unsafety. A broad understanding and definition of safety, allowing for instant measuring of unsafety. *Chem. Eng. Trans., 75*. https://doi.org/10.3303/CET1977043

Bocklisch, F., Paczkowski, G., Zimmermann, S., & Lampke, T. (2022). Integrating human cognition in cyber-physical Systems: A multidimensional fuzzy pattern model is applied to thermal spraying. *Journal of Manufacturing Systems, 63*, 162-176. https://doi.org/10.1016/j.jmsy.2022.03.005

Bolbot, V., Theotokatos, G., Bujorianu, L. M., Boulougouris, E., & Vassalos, D. (2019). Vulnerabilities and safety Assurance methods in Cyber-Physical Systems: A comprehensive review. *Reliability Engineering & System Safety 182*, 179-193. https://doi.org/10.1016/j.ress.2018.09.004

Broum, T., & Šimon, M. (2020). Safety requirements related to collaborative robots in the Czech Republic. *MM Science Journal*. https://doi.org/10.17973/MMSJ.2020_03_2019136

Burggraaf, J., Groeneweg, J., Sillem, S., Gelder, P. van (2019). How cognitive biases influence the data verification of safety indicators: a case study in rail. *Safety, 5*(4), 69. https://doi.org/10.3390/safety5040069

Byrne, E., Daykin, N., & Coad, J. (2016). Participatory Photography in Qualitative Research: A Methodological Review. *Visual Methodologies, 4*(2), 1–12.

Byrne, J.A. (2016). Improving the Peer Review of Narrative Literature Reviews. *Research Integrity and Peer Review, 1,* 12, 1–4. https://doi.org/10.1186/s41073-016-0019-2

Cabour, G., Ledoux, É., & Bassetto, S. (2021). A work-centered approach for cyber-physical-social system design: applications in aerospace industrial inspection. arXiv preprint arXiv:2101.05385.

Calvetti, D., Mêda, P., Chichorro Gonçalves M., & Sousa, H. (2020). Worker 4.0: The future of sensored construction sites. *Buildings, 10*, 169. https://doi.org/10.3390/buildings10100169

Cantelmi, R., Di Gravio, G., & Patriarca, R. (2021). Reviewing qualitative research approaches in the context of critical infrastructure resilience. *Environment Systems and Decisions, 41*(3), 341-376. https://doi.org/10.1007/s10669-020-09795-8

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. *Systems Engineering, 23*(2), 189-210. https://doi.org/10.1002/sys.21509

Challen, R., Denny, J., Pitt, M., Gompels, L., Edwards, T., & Tsaneva-Atanasova, K. (2019). Artificial intelligence, bias and clinical safety. *BMJ Quality & Safety, 28*(3), 231-237. https://doi.org/10.1136/bmjqs-2018-008370

Chaminda H. (2021). Opportunities, Challenges and Strategies for Integrating Cyber Security and Safety in Engineering Practice. *Engineering Technology Open Access Journal. 3*(5), 00119-00123. https://doi.org/10.19080/ETOAJ.2021.03.555622

Chehri, A., Fofana, I., & Yang, X. (2021). Security risk modeling in smart grid critical infrastructures in the era of big data and artificial intelligence. Sustainability, 13(6), 3196. https://doi.org/10.3390/su13063196

Cho, A. (2020). What Kinds of Sensors are Embedded in Smartphones? Newsroom. https://www.samsungsds.com/en/insights/What-Kinds-of-Sensors-are-Embedded-in-Smartphones.html [Accessed April 6, 2022]

Coscia E., Alonso R. & Soldatos J. (2021). A Review of Industrial Standards for AI in Manufacturing. Chapter 10 pp173-190. In John Soldatos and Dimosthenis Kyriazis (eds.). (2021). *Trusted Artificial Intelligence in Manufacturing: A Review of the Emerging Wave of Ethical and Human Centric AI Technologies for Smart Production*. Boston–Delft: Now Publishers. https://doi.org/10.1561/9781680838770.ch10

Costantino, F., Falegnami, A., Fedele, L., Bernabei, M., Stabile, S., & Bentivenga, R. (2021). New and Emerging Hazards for Health and Safety within Digitalized Manufacturing Systems. *Sustainability, 13*(19), 10948. https://doi.org/10.3390/su131910948

Cronin, P., Ryan, F., & Coughlan, M. (2008). Undertaking a Literature Review: A Step-by-Step Approach. *Br. J. Nurs. 2008*(17), 38–43. https://doi.org/10.12968/bjon.2008.17.1.28059

Deschacht, N. (2021). The digital revolution and the labour economics of automation: A review. *ROBONOMICS: The Journal of the Automated Economy, 1*, 8-8. https://journal.robonomics.science/index.php/rj/article/view/8

Dey, S., & Lee, S. W. (2021). Multilayered review of safety approaches for machine learning-based systems in the days of AI. *Journal of Systems and Software, 176*, 110941. https://doi.org/10.1016/j.jss.2021.110941

ESREDA (2015). Barriers to learning from incidents and accidents. ESReDA guidelines. https://www.esreda.org/wp-content/uploads/2021/01/ESReDA-barriers-learning-accidents-1.pdf

European Commission (2012). Directive 2012/18/EU of the European Parliament and of the Council of 4 July 2012 on Control of Major-Accident Hazards Involving Dangerous Substances (Seveso-III Directive). Available online: http://data.europa.eu/eli/dir/2012/18/oj [accessed on April, 12, 2020].

European Commission (2020a). Whitepaper on artificial intelligence –a European approach to excellence and trust. https://ec.europa.eu/info/sites/default/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf [Accessed May 25, 2022]

European Commission (2020b). Report on the safety and liability implications of artificial intelligence, the internet of things and robotics. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52020DC0064&from=en [Accessed may 25, 2022]

European Commission (2023). EU - General Product Safety Regulation (EU) 2023/988 (GPSR).

Fletcher S., & Webb P. (2017). Industrial Robot Ethics: The Challenges of Closer Human Collaboration in Future Manufacturing Systems, Chapter in: *A world with robots: International Conference on Robot Ethics: ICRE 2015*, Springer. https://doi.org/10.1007/978-3-319-46667-5_12

Forcina, A., & Falcone, D. (2021). The role of Industry 4.0 enabling technologies for safety management: A systematic literature review. *Procedia computer science, 180*, 436-445. https://doi.org/10.1016/j.procs.2021.01.260

Fosch-Villaronga, E., & Mahler, T. (2021). Cybersecurity, safety and robots: Strengthening the link between cybersecurity and safety, in the context of care robots. *Computer Law and Security Review, 41*, 105528. https://doi.org/10.1016/j.clsr.2021.105528

Furstenau L.B., Rodrigues Y.P.R., Sott M.K., Leivas P., Dohan M.S., López-Robles J.R., Cobo M.J., Bragazzi N.L., & Choo K.K.R. (2022). Internet of things: Conceptual network structure, main challenges and future directions. *Digital Communications and Networks, 9(*3), 677-687. https://doi.org/10.1016/j.dcan.2022.04.027

Garibaldo, F., & Rebecchi, E. (2018). Cyber-physical system. *AI & Society, 33*(3), 299-311. https://doi.org/10.1007/s00146-018-0802-3

Gharib, M., Lollini, P., & Bondavalli, A. (2021). IQCPSoS: A Model-Based Approach for Modeling and Analyzing Information Quality Requirements for Cyber-Physical System-of-Systems. *Journal of Data Semantics 10*(3): 267 289. https://doi.org/10.1007/s13740-021-00129-8

Girs, S., Sentilles, S., Asadollah, S. A., Ashjaei, M., & Mubeen, S. (2020). A systematic literature study on definition and modeling of service-level agreements for cloud services in IoT. *IEEE Access, 8*, 134498-134513. https://doi.org/10.1109/ACCESS.2020.3011483

Gjerdrum, D., & Peter, M. (2011). The new international standard on the practice of risk management – a comparison of ISO 31000:2009 and the COSO erm framework. *Risk Management, 21*, 8–12.

Google White Paper (2019). Perspectives on issues in AI governance.; some of the safety considerations to ensure safety

are rightly identified by Google

Gray J., Liu D.T., Nieto-Santisteban M., Szalay A.S., DeWitt D., & Heber G. (2005). Scientific Data Management in the Coming Decade. *Association for Computing Machinery, 34*(4), 34-41. https://doi.org/10.1145/1107499.1107503

Gromov, G. (1995). Roads and Crossroads of the Internet History. http://www.netvalley.com/cgi-bin/intval/net_history.pl?chapter=1 [Accessed March 29, 2022]

Gualtieri, L., Rauch, E., Rojas, R., Vidoni, R., & Matt, D. T. (2018). Application of Axiomatic Design for the design of a safe collaborative human-robot assembly workplace. *MATEC Web of Conferences, 223*, 01003. EDP Sciences. https://doi.org/10.1051/matecconf/201822301003

Gülen, K. (2021). 1nm processors to be manufactured by 2030. https://techbriefly.com/2021/10/01/asml-predicts-1-nm-processors-to-be-manufactured-by-2030/ [Accessed March 30, 2022]

Haidegger, T., Virk, G. S., Herman, C., Bostelman, R., Galambos, P., Györök, G., & Rudas, I. J. (2020). Industrial and Medical cyber-physical systems: Tackling user requirements and challenges in robotics. In *Recent advances in intelligent engineering* (pp. 253-277). Springer, Cham. https://doi.org/10.1007/978-3-030-14350-3_13

Hamon, R., Junklewitz, H., & Sanchez, I. (2020). Robustness and explainability of artificial intelligence. *Publications Office of the European Union.* https://doi.org/10.2760/57493

Hopcraft, R., Tam, K., Moara_Nkwe, K., & Jones, K. (2021). The Development of a Cyber Safety Culture. ErgoSHIP 2021, 2nd-3rd September 2021.

Huang L., Wu C., & Wang B. (2019). Challenges, opportunities and paradigm of applying big data to production safety Management: From a theoretical perspective. *Journal of Cleaner Production, 231*, 592-599. https://doi.org/10.1016/j.jclepro.2019.05.245

Huang L., Wu C., Wang B., & Ouyang Q. (2018). Big-data-driven safety decision-making: A conceptual framework and its influencing factors. *Safety Science, 109*, 46-56. https://doi.org/10.1016/j.ssci.2018.05.012

Huang, Z., Jiang, X., Chen, L., & Fan, D. (2019). Research on safe communication architecture for real-time Ethernet distributed control system. *IEEE Access, 7*, 89821-89832. https://doi.org/10.1109/ACCESS.2019.2926650

Huff D. (1991). *How to lie with statistics.* Penguin books. ISBN. 9780140136296.

IEEE (2016). Autonomous and Intelligence Systems. April 2016. https://ethicsinaction.ieee.org/

Jefroy, N., Azarian, M., & Yu, H. (2022). Moving from Industry 4.0 to Industry 5.0: What Are the Implications for Smart Logistics? *Logistics* 2022, *6*, 26. https://doi.org/10.3390/logistics6020026

Jiang, Y., Atif, Y., Ding, J., & Wang, W. (2019). A Semantic Framework with Humans in the Loop for Vulnerability-Assessment in Cyber-Physical Production Systems. In International Conference on Risks and Security of Internet and Systems, October 2019. (pp. 128-143). *Springer*, Cham.

Jing Q., A. V. & Vasilakos et al. (2014). Security of the internet of things: perspectives and challenges. *Wireless Networks, 20*(8), 2481–2501. https://doi.org/10.1007/s11276-014-0761-7

Johnsen S., Skramstad T., & Hagen J. (2009). Enhancing the safety, security and resilience of ICT and SCADA systems using action research. *IFIP Advances in Information and Communication Technology, 311*, 113–124. http://doi.org/10.1007/978-3-642-04798-5_8

Khalid, A., Kirisci, P., Ghrairi, Z., Pannek, J., & Thoben, K. D. (2017). Safety requirements in collaborative human–robot cyber-physical system. In Dynamics in logistics (pp. 41-51). Springer. https://doi.org/10.1007/978-3-319-45117-6_4

Khalid, A., Kirisci, P., Khan, Z. H., Ghrairi, Z., Thoben, K. D., & Pannek, J. (2018). Security framework for industrial Collaborative robotic cyber-physical system. *Computers in Industry, 97*, 132-145. http://doi.org/10.1016/j.compind.2018.02.009

Khan F., Abunada H., John D. & Enmesh T. (2009). Development of Risk-Based Process Safety Indicators. *Process Safety Progress, 29*(2), 133-143. https://doi.org/10.1002/prs.10354

Kirwan, B. (2008). From safety culture to safety intelligence, p 18-23. In: Kao, T. (Ed.), Probabilistic Safety Assessment and Management. *PSAM9.*

Kline R.R. (2011). Cybernetics, Automata Studies, and the Dartmouth Conference on Artificial Intelligence. The Dartmouth Summer Research Project on Artificial Intelligence. *IEEE Annals of the History of Computing, 33*(4). https://doi.org/10.1109/MAHC.2010.44

Koene, A., Smith, A. L., Egawa, T., Mandalh, S., & Hatada, Y. (2018). IEEE P70xx, Establishing standards for ethical technology. Proceedings of KDD, ExCeL London UK, August, 2018 (KDD'18).

Laurent, A., & Fabiano, B. (2022). A Critical Perspective on the Impact of Industry 4.0's New Professional Safety Management Skills on Process Safety Education. *Chemical Engineering Transactions, 91*, 67-72. https://doi.org/10.3303/CET2291012

Lee, T. (2019). The global rise of "fake news" and the threat to democratic elections in the USA. *Public Administration and Policy, 22*(1), 15-24. https://doi.org/10.1108/PAP-04-2019-0008

Leiber, T. (2017). Computational social science and big data: a quick SWOT analysis. In *Berechenbarkeit der Welt?* Springer VS, Wiesbaden. https://doi.org/10.1007/978-3-658-12153-2_14

Leong, C. W. K. (2018). Managing Epistemic Uncertainties in the Underlying Models of Safety Assessment for Safety-Critical Systems (Doctoral dissertation, University of York).

Leslie, D. (2019). Understanding artificial intelligence ethics and safety: A guide for the responsible design and

implementation of AI systems in the public sector. https://doi.org/10.2139/ssrn.3403301f
Leveson, N.G., & Stephanopoulos, G. (2014). A system-theoretic, control-inspired view and approach to process safety. *AiChE Journal, 60*(1). https://doi.org/10.1002/aic.14278
Li, Y., & Guldenmund, F.W. (2018). Safety management systems: a broad overview of the literature. *Safety and Security Science, 103*, 94–123. https://doi.org/10.1016/j.ssci.2017.11.016
Lindhout P., Kingston J,C. & Reniers, G. (2019). Learning from language problem related accident information in the process industry: A literature study. *Process Safety and Environmental Protection*, *130,* 140–152. https://doi.org/10.1016/j.psep.2019.06.017
Lindhout P., Kingston J,C., Hansen F.T., & Reniers G. (2020). Reducing unknown risk: The safety engineers' new horizon. *Journal of Loss Prevention in the Process Industries, 68,* 104330. https://doi.org/10.1016/j.jlp.2020.104330
Liu, Y. (2022). Risk management of smart healthcare systems: Delimitation, state-of-arts, process, and perspectives. *Journal of Patient Safety and Risk Management.* https://doi.org/10.1177/25160435221102242
Mageto, J. (2021). Big data analytics in sustainable supply chain management: A focus on manufacturing supply chains. *Sustainability, 13*(13), 7101. https://doi.org/10.3390/su13137101
Magrabi, F., Ammenwerth, E., McNair, J. B., De Keizer, N. F., Hyppönen, H., Nykänen, P., ... & Georgiou, A. (2019). Artificial intelligence in clinical decision support: Challenges in evaluating AI and practical implications. *Yearbook of medical informatics, 28*(1), 128-134. https://doi.org/10.1055/s-0039-1677903
Malhotra, P., Singh, Y., Anand, P., Bangotra, D. K., Singh, P. K., & Hong, W. C. (2021). Internet of things: Evolution, concerns and security challenges. Sensors, 21(5), 1809. https://doi.org/10.3390/s21051809
Marjumin, N. H., Sidek, S., Hassan, M. A., Rajikon, M., & Kamalrudin, M. (2019). The challenges and contribution of internet of things (Iot) for smart living. International. *Journal of Recent Technol. and Eng., 8,* 162-166.
Matta, G., Chlup, S., Shaaban, A. M., Schmittner, C., Pinzenöhler, A., Szalai, E., & Tauber, M. (2021). Risk Management and Standard Compliance for Cyber-Physical Systems of Systems. *Infocommunications Journal, 13*(2), 32-39. https://doi.org/10.36244/ICJ.2021.2.5
McKee, D. W., Clement, S. J., Almutairi, J., & Xu, J. (2017). Massive-scale automation in cyber-physical systems: Vision & challenges. In March 2017 IEEE 13th International Symposium on Autonomous Decentralized System (ISADS) (pp. 5-11). *IEEE*. https://doi.org/10.1109/ISADS.2017.56
Medina, A. C., Mora, J. F., Martinez, C., Barrero, N., & Hernandez, W. (2019). Safety protocol for collaborative human-robot recycling tasks. *IFAC-PapersOnLine, 52*(13), 2008-2013. https://doi.org/10.1016/j.ifacol.2019.11.498
Mezzanotte, P. (2019). Industrial Internet of Things: Challenges, Opportunities, and Directions. *IEEE Transactions on Industrial Informatics, 14.* https://doi.org/10.1109/TII.2018.2852491
Moencks, M., Roth, E., Bohné, T., & Kristensson, P. O. (2021). Augmented Workforce: Contextual, Cross-hierarchical Enquiries on Human-technology Integration in Industry. *Computers & Industrial Engineering*, *165*, 107822. https://doi.org/10.1016/j.cie.2021.107822
Mohammadpoor, M., & Torabi, F. (2020). Big Data analytics in oil and gas industry: An emerging trend. *Petroleum, 6*(4), 321-328. https://doi.org/10.1016/j.petlm.2018.11.001
Mol, T. (2003). *Productive Safety Management*. Butterworth-Heinemann, Oxford, UK. https://doi.org/10.4324/9780080474021
Moore G.E. (1965). Cramming more components onto integrated circuits. *Electronics, 38*(8) 1965.
Noah Jr, P.D. (2017). A Systematic Approach to the Qualitative Meta-Synthesis. *Issues Inf. Syst. 18*(2), 196-205. https://doi.org/10.48009/2_iis_2017_196-205
O'Neil, C. (2017). *Weapons of math destruction: how big data increases inequality and threatens democracy*. Crown Publishing Group
Ogbuke, N. J., Yusuf, Y. Y., Dharma, K., & Mercangoz, B. A. (2022). Big data supply chain analytics: ethical, privacy and the security challenges posed to businesses, industries, and society. *Production Planning and Control, 33*(2-3), 123-137. https://doi.org/10.1080/09537287.2020.1810764
Olowononi, F. O., Rawat, D. B., & Liu, C. (2020). Resilient machine learning for networked cyber physical systems: A survey for machine learning security to securing machine learning for cps. *IEEE Communications Surveys & Tutorials, 23*(1), 524-552. https://doi.org/10.1109/COMST.2020.3036778
Olszewska, J. I., Houghtaling, M., Goncalves, P. J., Fabiano, N., Haidegger, T., Carbonera, J. L., ... & Prestes, E. (2020). Robotic standard development life cycle in action. *Journal of Intelligent and Robotic Systems 98*(1), 119-131. https://doi.org/10.1007/s10846-019-01107-w
Pandita, R., & Singh, S. (2011). Grey literature: A valuable untapped stockpile of information. *Journal of the Young Librarians Association, 5*, 52–60.
Pasman H., & Rogers, W. (2014). How can we use the information provided by process safety performance indicators? Possibilities and limitations. *Journal of Loss Prevention in the Process Industries, 30*, 197-206. https://doi.org/10.1016/j.jlp.2013.06.001
Patriarca R., Di Gravio G., Cioponea R., & Licu A. (2019). Safety intelligence: Incremental proactive risk management for holistic aviation safety performance. *Safety Science, 118*, 551-567. https://doi.org/10.1016/j.ssci.2019.05.040
Patriarca, R., Falegnami, A., Costantino, F., Di Gravio, G., De Nicola, A., & Villani, M. L. (2021). WAx: An integrated

conceptual framework for the analysis of cyber-socio-technical systems. *Safety Science, 136*, 105142. https://doi.org/10.1016/j.ssci.2020.105142

Pawson, R., Wong, G., Lesley, & Owen L. (2011). Known knowns, known unknowns, unknown unknowns: the predicament of evidence-based policy. *Am. J. Eval. 32* (4), 518–546. https://doi.org/10.1177/1098214011403831

Peña-Casas, R., Ghailani, D., & Coster, S. (2018). The impact of digitalization on job quality in European public services. Homecare and employment of service workers. *European Social Observatory, European Public Services. Union.* Available at: http://epsu.org/sites/default/files/article/files/FINAL%20REPORT%20EPSU%20DIGITALISATION%20-%20OSE%20June%202018.pdf

Pereira, A., & Thomas, C. (2020). Challenges of machine learning applied to safety-critical cyber-physical systems. *Machine Learning and Knowledge Extraction, 2*(4), 579-602. https://doi.org/10.3390/make2040031

Peserico, G., Morato, A., Tramarin, F., & Vitturi, S. (2021). Functional Safety Networks and Protocols in the Industrial Internet of Things Era. *Sensors, 21*(18), 6073. https://doi.org/10.3390/s21186073

Pillay, M. (2015). Accident causation, prevention and safety management: a review of the state-of-the-art. *Procedia Manufact, 3*, 1838–1845. https://doi.org/10.1016/j.promfg.2015.07.224

Pogliani, M., Quarta, D., Polino, M., Vittone, M., Maggi, F., & Zanero, S. (2019). Security of controlled manufacturing systems in a connected factory for industrial robots. *Journal of Computer Virology and Hacking Techniques 15*(3), 161-175. https://doi.org/10.1007/s11416-019-00329-8

Purcell, B. (2012). The emergence of "big data" technology and analytics. *Journal of technology research, 4*(1), 1-7.

Putnik, G. D., Ferreira, L., Lopes, N., & Putnik, Z. (2019). What is a Cyber-Physical System: Definitions and models spectrum. Fme Transactions, 47(4), 663-674. http://doi.org/10.5937/fmet1904663P

Ram J., Afridi N.K., & Khan K.A. (2019). "Adoption of Big Data analytics in construction: development of a conceptual model". *Built Environment Project and Asset Management, 9*(4). https://doi.org/10.1108/BEPAM-05-2018-0077

Renaud K., Zimmermann V., Schürmann T., & Böhm C. (2021). Exploring cybersecurity-related emotions and finding that it is challenging to measure. *Humanities and Social Sciences Communications 8*(75), 1-18. https://doi.org/10.1057/s41599-021-00746-5

Reniers G., & Amyotte P. (2012). Prevention in the chemical and process industries: Future directions. *Journal of Loss Prevention in the Process Industries 25*(1), 227-231. https://doi.org/10.1016/j.jlp.2011.06.016

Reniers, G., & Khakzad, N. (2017). Revolutionizing safety and security in the chemical and process industry: applying the CHESS concept. *Journal of Integrated Security and Safety Science*, *1*(1). https://doi.org/10.18757/jiss.2017.1.1547

Riordan M. (1997). The Incredible Shrinking Transistor. https://www.technologyreview.com/1997/11/01/237128/the-incredible-shrinking-transistor/ [Accessed March 30, 2022]

Roser M. & Ritchie H. (2020). Transistor count over time. https://commons.wikimedia.org/w/index.php?curid=98219918

Salvini P., Paez-Granados D., & Billard A. (2021). On the Safety of Mobile Robots Serving in Public Spaces: Identifying gaps in EN ISO 13482:2014 and calling for a new standard. *ACM Trans. Hum.-Robot Interact., Vol. 10*, No. 3, Article 19. https://doi.org/10.1145/3442678

Schoitsch, E., & Schmittner, C. (2020). Ongoing cybersecurity and safety standardization activities related to highly automated/autonomous vehicles. *In Intelligent System Solutions for Auto Mobility and Beyond*. Springer, Cham. https://doi.org/10.1007/978-3-030-65871-7_6

Schwab, K. (2015). The Fourth Industrial Revolution: What It Means and How to Respond. https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

Sculley D., Holt G., Golovin D., Davydov E., Phillips T., Ebner D., Chaudhary V., Young M., Crespo J.-F., & Dennison D. (2015). Hidden technical debt in machine learning systems. *Advances in neural information processing systems, 2*, 2503-2511.

See, J. E., Drury, C. G., Speed, A., Williams, A., & Khalandi, N. (2017). The Role of Visual Inspection in the 21 Century. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 61*(1), 262-266. https://doi.org/10.1177/1541931213601548

Shahriari, K., Shahriari, M. (2017). Ethically Aligned Design: A Vision for Prioritizing Human Wellbeing with Autonomous and Intelligent Systems technical report. *IEEE*, https://doi.org/10.1109/IHTC.2017.8058187

Shneiderman, B. (2020). Human-centered artificial intelligence: Reliable, safe & trustworthy. *International Journal of Human–Computer Interaction, 36*(6), 495-504. https://doi.org/10.1080/10447318.2020.1741118

Smith, S.K., Mountain, G.A., & Hawkins, R.J. (2015). A scoping review to identify the techniques frequently used when analysis of qualitative visual data. *International Journal of Social Research Methodology, 19* (6). pp. 693-715. https://doi.org/10.1080/13645579.2015.1087141

Soldatos, J., & Kyriazis, D. (2021). *Trusted Artificial Intelligence in Manufacturing: A Review of the Emerging Wave of Ethical and Human Centric AI Technologies for Smart Production*. nowPublishers. https://doi.org/10.1561/9781680838770

Spiekermann, S., & Winkler, T. (2020). Value-based engineering for ethics by design. https://doi.org/10.48550/arXiv.2004.13676

Stefana, E., & Paltrinieri, N. (2021). ProMetaUS: a proactive meta-learning uncertainty-based framework to select models for Dynamic Risk Management. *Safety science, 138*, 105238. https://doi.org/10.1016/j.ssci.2021.105238

Stenson R. (2016). Is This the First Time Anyone Printed, 'Garbage In, Garbage Out'? https://www.atlasobscura.com/articles/is-this-the-first-time-anyone-printed-garbage-in-garbage-out [Accessed April 4, 2022]

Sun, S., Zheng, X., Gong, B., Garcia Paredes, J., & Ordieres-Meré, J. (2020). Healthy operator 4.0: A human cyber–physical system architecture for smart workplaces. *Sensors, 20*(7). https://doi.org/10.3390/s20072011

Sun, X.; Yu, H.; Solvang, W.; Wang, Y.; & Wang, K. (2021). The application of Industry 4.0 technologies in sustainable logistics: A systematic literature review (2012–2020) to explore future research opportunities. *Environ. Sci. Pollut. Res. 29*, 9560–9591. https://doi.org/10.1007/s11356-021-17693-y

Swuste P., Theunissen J., Schmitz P., Reniers G., & Blokland P. (2016). Process safety indicators : a review of literature *Journal of loss prevention in the process industries 40*(2016),162-173. https://doi.org/10.1016/j.jlp.2015.12.020

Timulak, L. (2014). Qualitative meta-analysis. In: Flick, U. (Ed.) (2014) Mapping the field. *The SAGE Handbook of Qualitative Data Analys*is, Pp. 481-495. London UK: Sage Publications Ltd.

Tramarin, F., Mok, A. K., & Han, S. (2019). Real-time and reliable industrial control over wireless lans: Algorithms, protocols, and future directions. *Proceedings of the IEEE, 107*(6), 1027-1052. https://doi.org/10.1109/JPROC.2019.2913450

Trentesaux, D., & Karnouskos, S. (2019, October). Ethical behaviour aspects of autonomous intelligent cyber-physical systems. In *Service Oriented, Holonic and Multi-agent Manufacturing Systems for Industry of the Future.* Springer, Cham. https://doi.org/10.1007/978-3-030-27477-1_5

Tschider, C. A. (2018). Regulating the internet of things: discrimination, privacy, and cybersecurity in the artificial intelligence age. DENv. L. REv., 96, 87. https://doi.org/10.2139/ssrn.3129557

Uraikul, V., Chan, C. W., & Tontiwachwuthikul, P. (2007). Artificial intelligence for monitoring and supervisory control of process systems. *Engineering applications of artificial intelligence, 20*(2), 115-131. https://doi.org/10.1016/j.engappai.2006.07.002

Van Acker, B. (2020). Mental Workload Monitoring in the Manufacturing Industry: Conceptualisation, Operationalisation and Implementation (Doctoral dissertation, Ghent University). http://hdl.handle.net/1854/LU-8682112

Vierendeels, G., Reniers, G., Nunen, K. van, & Ponnet, K. (2018). An integrative conceptual framework for safety culture: The Egg Aggregated Model (TEAM) of safety culture. *Safety science, 103*, 323-339. https://doi.org/10.1016/j.ssci.2017.12.021

Vinuesa, R., Azizpour, H., Leite, I., Balaam, M., Dignum, V., Domisch, S., ... & Fuso Nerini, F. (2020). The role of artificial to achieve Sustainable Development Goals. *Nature Communications, 11*(1), 1-10. https://doi.org/10.1038/s41467-019-14108-y

Vogt, J. (2021). Where is the human got to go? Artificial intelligence, machine learning, big data, digitalisation, and human–robot interaction in Industry 4.0 and 5.0. *AI & Soc, 36*, 1083–1087. https://doi.org/10.1007/s00146-020-01123-7

Wang B., Wang Y., Yan F., & Zhao W. (2022). Safety intelligence toward safety management in a big-data environment: A general model and its application in urban safety management. *Safety Science, 154.* https://doi.org/10.1016/j.ssci.2022.105840

Wang B., & Wang Y. (2021). Big data in safety management: An overview. *Safety Science, 143*. https://doi.org/10.1016/j.ssci.2021.105414

Wang, B. (2019). Demystifying safety-related intelligence in safety management: Some key questions answered from a theoretical perspective. *Safety Science, 120*(4). https://doi.org/10.1016/j.ssci.2019.08.030

Wang, B., Wu, C., Huang, L., & Kang, L. (2019). Using data-driven safety decision-making to realize smart safety management in the era of big data: A theoretical perspective on basic questions and their answers. *Journal of Cleaner Production, 210,* 1595–1604. https://doi.org/10.1016/j.jclepro.2018.11.181

Wang, J., Xu, C., Zhang, J., & Zhong, R. (2021). Big data analytics for intelligent manufacturing systems: A review. *Journal of Manufacturing Systems, 62*. https://doi.org/10.1016/j.jmsy.2021.03.005

Wessels, R. H. A. (1997). Het belang en toegankelijkheid van grijze literatuur. [The Importance and Accessibility of Grey Literature]. *Informatie Professional, 1*(3), 28-32.

Wirth, N. (1995). A plea for lean software. *Computer, 28*(2), 64-68. https://doi.org/10.1109/2.348001

Yilma, B. A., Panetto, H., & Naudet, Y. (2021). Systemic formalisation of Cyber-Physical-Social System (CPSS): A systematic literature review. *Computers in Industry, 129*, 103458. https://doi.org/10.1016/j.compind.2021.103458

Zheng, N., Liu, Z., Ren, P., Ma, Y., Chen, S., Yu, S., Xue, J., Chen, B., & Wang, F. (2017). Hybrid-augmented intelligence : Collaboration and cognition. *Frontiers of Information Technology & Electronic Engineering, 18*(2), 153-179. https://doi.org/10.1631/FITEE.1700053

Zimmermann, V., & Renaud, K. (2019). Moving from a 'human-as-problem" to a 'human-as-solution" cybersecurity mindset. *International Journal of Human-Computer Studies, 131*, 169-187. https://doi.org/10.1016/j.ijhcs.2019.05.005

APPENDIX 1: Standards relevant to CPS safety and security

ANSI/RIA R15.06; ANSI/RIA R15.06 Cobots safety standard North America

EN 62304:2015; EN 62304:2006 A1:2015. European Standard Medical device software –software life-cycle processes.

ETSI TR 103375; ETSI (2016). SmartM2M; IoT Standards landscape and future evolutions. Technical Report TR 103375. European Telecommunications Standards Institute (ETSI), Sophia Antipolis, France.

IEC 60601:2017; IEC 60601-4-1:2017 collateral standard (Medical electrical equipment—Part 4-1: Guidance and interpretation—Medical electrical equipment and medical electrical systems employing a degree of autonomy)

IEC 61508:2016; IEC 61508:2016 Functional Safety of Electrical/Electronic/Programmable Electronic Safety-Related Systems; Standard; International Electrotechnical Commission: Geneva, Switzerland.

IEC 62046:2018; IEC 62046:2018, Safety of machinery— Application of protective equipment to detect the presence of persons, Geneva, Switzerland: International Electrotechnical Commission, (2018).

IEC 62541; IEC 62541 International Electrotechnical Commission: 62541-1 Ed. 1.0: OPC Unified Architecture Specification - Part 1: Overview and Concepts. IEC Std., 2008.

IEEE 1872:2015; IEEE 1872–2015 Ontologies for Robotics and Automation (https://standards.ieee.org/develop/wg/ORA.html)

IEEE P1872.2; IEEE P1872.2 Standard for Autonomous Robotics

IEEE P7001:2017; IEEE (2017) P7001—Transparency of Autonomous Systems

IEEE P7000:2017; IEEE (2017) P7000—Model Process for Addressing Ethical Concerns During System Design

ISO 10075; ISO 10075 Ergonomic principles related to mental workload, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/66900.html (Accessed: June 2021).

ISO 10218-1,-2:2011; Robots and robotic devices – Safety requirements for industrial robots, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/51330.html (Accessed: June 2021).

ISO 12100:2021; ISO 12100 Safety of machinery – General principles for design – Risk assessment and risk reduction, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/51528.html (Accessed: June 2021).

ISO 13482:2014; ISO 13482:2014 Robots and robotic devices—Safety requirements for personal care robots. International Organization for Standardization. Geneva

ISO 13485:2003; ISO 13485:2003 Medical devices—Quality management systems— Requirements for regulatory purposes

ISO 13854:2017; ISO 13854:2017, Safety of machinery— Minimum gaps to avoid crushing of parts of the human body, Geneva, Switzerland: International Organization for Standardization, (2017).

ISO 13855:2010; ISO 13855 (Safety of machinery—Positioning of safeguards with respect to the approach speeds of parts of the human body, 2002),2010

ISO 13857:2008; ISO 13857:2017, Safety of machinery— Safety distances to prevent hazard zones being reached by upper and lower limbs, Geneva, Switzerland: International Organization for Standardization, (2017).

ISO 14001; ISO 14000 family – Environmental management, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/iso-14001-environmental-management.html (Accessed: June 2021).

ISO 14118:2015; ISO, 14118:2015, Safety of machinery— Prevention of unexpected start-up, Geneva, Switzerland: International Organization for Standardization.

ISO 14120:2015; ISO, 14120: 2015, Safety of machinery—Guards—General requirements for the design and construction of fixed and movable guards, Geneva, Switzerland: International Organization for Standardization, (2015).

ISO 26000; ISO 26000 Social responsibility, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/iso-26000-social-responsibility.html (Accessed: June 2021).

ISO 26262:2018-12; ISO 26262:2018-12: Road Vehicles—Functional Safety; International Standards Organization: Geneva, Switzerland, 2018.

ISO 26800; ISO 26800 Ergonomics – General approach, principles and concepts, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/42885.html (Accessed: June 2021).

ISO 31000; ISO 31000 Risk management, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/iso-31000-risk-management.html (Accessed: June 2021).

ISO 45001; ISO 45001 Occupational health and safety, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/iso-45001-occupational-health-and-safety.html (Accessed: June 2021).

ISO 6385; ISO 6385 Ergonomics principles in the design of work systems, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/63785.html (Accessed: June 2021).

ISO 8373:2012; ISO 8373:2012 Robots and robotic devices –vocabulary.

ISO 9000; ISO 9000 family – Quality management, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/iso-9001-quality-management.html (Accessed: June 2021).

ISO 9241; ISO 9241 Ergonomics of human-system interaction, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/77520.html (Accessed: June 2021).

ISO/CD 8373:2019; ISO/CD 8373:2019 Industrial robot, International Organization for Standardization

ISO/DTR 23482-1; ISO/DTR 23482-1—Technical report: Validation criteria for personal care robots

ISO/IEC 18033; ISO/IEC 18033 Information technology – Security techniques – Encryption algorithms, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/76156.html (Accessed: June 2021).

ISO/IEC 27001; ISO/IEC 27001 Information Security Management, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/isoiec-2

ISO/IEC 27002; ISO/IEC 27002 Information technology – Security techniques – Code of practice for information security controls, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/545 33.html (Accessed: June 2021).

ISO/IEC 27040; ISO/IEC 27040 Information technology – Security techniques – Storage security, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/44404.html (Accessed: June 2021).

ISO 27701:2019 Security techniques. Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management. Requirements and guidelines (extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy management within the context of the organization)

ISO/IEC 29100; ISO/IEC 29100 Information technology – Security techniques – Privacy framework. International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/45123.html (Accessed: June 2021).

ISO/PRF TR 23482-2; ISO/PRF TR 23482-2—Application guide for ISO 13482, Part 2: Application Guide

ISO/TR 16982; ISO/TR 16982 Ergonomics of human-system interaction – Usability methods supporting human-centred design, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/31176.html (Accessed: June 2021).

ISO/TS 13849:2018; ISO/TS 15066 Robots and robotic devices – Collaborative robots, International Organization for Standardization, Geneva, Switzerland. https://www.iso.org/standard/62996.html (Accessed: June 2021).

ISO/TS 15066:2016; ISO/TS13849 (2018). ISO - International Organization for Standardization. URL https://www.iso.org/standard/69883.html.

SAE ARP4754A:2010 SAE International. ARP4754A:2010: Guidelines for Development of Civil Aircraft and Systems; SAE International: Warrendale, PA, USA, 2010.