Dipartimento di Ingegneria Informatica Automatica e Gestionale
Facoltà di Ingegneria dell'Informazione, Informatica e Statistica

Dipartimento di Ingegneria Meccanica ed Aereospaziale
Facoltà di Ingegneria Civile e Industriale

Tesi di Dottorato

HUMANS IN CYBER RESILIENCE: MANAGERIAL AND OPERATIONAL OPPORTUNITIES

Dottoranda
Silvia Colabianchi

Tutor
Francesco Costantino
Fabio Nonino

Dottorato di Ricerca in Ingegneria Industriale e Gestionale
Ciclo XXXV – 2019-2022

HUMANS IN CYBER RESILIENCE:
MANAGERIAL AND OPERATIONAL OPPORTUNITIES

Summary

The hyper-connected environment of today has resulted in a substantial boost in productivity, efficiency, and system integration, but it has also raised the number of possible threats. Organizations are increasingly reliant on data and information from their interconnected systems, making them exposed to a variety of cyber attacks. Cyber threats have an impact on the continuity of their company operations, the loss of confidential information, reputational harm, and possibly the safety of their employees. Cyber attacks have become more sophisticated, raising awareness of the importance of not limiting the design of cybersecurity practices to detection and protection phases, but of considering the ability to respond, recover, and thus withstand cyber incidents as fundamental from a cyber resilience perspective.

This thesis is based on four comprehensive research objectives. First, the thesis sheds light on the definitions and topics related to cyber resilience and cyber security. These analyses set the groundwork and motivate the challenges addressed in the thesis. The second part of the thesis then focuses on the need to go beyond purely technical aspects when managing cyber resilience by integrating organizational and human aspects. The debate is developing as to which is the human role in cyber socio-technical systems. Specifically, the aim is to identify new managerial and operational opportunities to raise the positive role of humans in increasing the cyber resilience of the cyber socio-technical systems in which organizations operate today.

The thesis maps the human factors involved in cybersecurity, identifying under what circumstances they can be a driver or a barrier to it, helping practitioners prioritize and achieve cyber resilience goals. Second, tools that can be used as external leverage to improve human integration with cyber socio-technical systems are presented. Outsourcing strategies for cybersecurity management are discussed. In addition, a reference architecture and taxonomy for intelligent digital assistants is developed and a proactive agent to support employees in managing cybersecurity issues is tested.

In sum, this thesis adds both theoretical and practical contributions to the field of cyber resilience, focusing on managerial and operational opportunities. The thesis has a publication-based structure.

Table of Contents

1. Introduction

The widespread adoption of cyber-physical systems, the Internet of Things (IoT), big data, smart technologies, and cloud computing across industries has led to a growing need to develop information security systems. Working with large groups of devices, software, and interconnected systems can cause some of them to be compromised. Moreover, the presence of these sensors in industrial environments poses a considerable security challenge since most systems were not designed with cybersecurity in mind, and people involved in these processes are often not trained to face these new emerging challenges. These new hazards are constantly evolving, forcing a continuous rethinking of strategies to ensure business continuity. Moreover, threats and incidents have become more sophisticated, increasing awareness of not limiting the design of cybersecurity practices to the phases of detection and protection, but of considering as fundamental the ability to respond, recover and therefore withstand cyber incidents in a cyber resilience perspective. In this scenario, organizations find themselves working in a more complex environment and the theory suggested to cope with complexity is Resilience Engineering which addresses ways to build an adaptive capacity to cope with intractable systems. Specifically, related to the topic of cyber threats and new hazards that shape today's cyberspace, the concept of cyber resilience has become popular in recent years.

The first step of this thesis work was to harmonize knowledge on the topic of cybersecurity and cyber resilience. For this reason, the first research question was addressed.

- RQ1: Which are the approaches to enhance cyber resilience?

To answer the question a systematic literature review exploring the topics of cybersecurity, cyber resilience, and cyber-socio-technical systems was conducted. The study of the literature made it possible to compare definitions of cybersecurity and cyber resilience, collect and describe the main cyber threats and attacks, and present internationally recognized frameworks for managing cyber resilience in the enterprise. The author will define cyber resilience as *"the ability to continuously deliver the intended outcome despite adverse cyber events caused by humans and nature"*[1] throughout the thesis. Cyber resilience it is an ongoing practice, not a one-time effort. It is the ability to continually adapt to new or potential risks. In this scenario during the first stage of the thesis difference between cybersecurity and cyber resilience are underlined, emphasizing the concept that the main distinction between cybersecurity and cyber resilience comes in the objective and scope. The purpose of cybersecurity is to keep IT systems operational, whereas cyber resilience during adverse events focuses on maintaining business objectives.

This initial analysis showed how the literature on the topic of cyber resilience is still quite broad and non-domain specific and lacks a commonly endorsed definition. However, when placed in relation to the better-known topic of resilience engineering, the need for adaptive capacity during cyber incidents and the focus on organizational aspects and human perspective over purely technical aspects of cybersecurity are highlighted. In this context, the debate is developing as to which is the human role in cyber socio-technical systems. Specifically, the increasing complexity and interconnectedness of these systems are forcing researchers to investigate in which part humans are a threat or an opportunity for cyber resilience. This laid the foundation for the research direction followed in the second part of the thesis.

The second step of the thesis identified new managerial and operational opportunities to raise the positive role of humans in increasing the cyber resilience of cyber socio-technical systems in which organizations operate today. To follow this research direction, three additional research questions were defined.

- RQ2: A human-centric cyber resilience
    - RQ2.1: Which are the human factors involved in cybersecurity?
    - RQ2.2**:** How does each factor contribute as a weakness or opportunity to cyber resilience?

Studies addressed humans as flexible and able to rapidly judge and attack, stressing the importance of continuous cooperation between humans and machines to pursue cyber resilience effectiveness. However, recognizing the centrality of the individual in the system yet not knowing his or her potential and vulnerabilities does not achieve a successful organizational cyber resilience strategy. Therefore, to answer the question an in-depth analysis of the human factor element involved in cybersecurity was conducted. The purpose was to identify all the factors involved which may influence individuals and generate both positive and negative actions on the system. Each factor has been linked with the available reference and clear motivations for the identified relationship explained. Then, from a more qualitative point of view, the integration of the NIST cybersecurity framework with cybersecurity-related human factors has been investigated. The NIST framework provides a set of standards, guidelines, and best practices for managing cybersecurity risks and enhancing cyber resilience. Although the framework provides outcome-oriented statements that give considerations for creating or improving a cyber resilience program, it lacks best practices to leverage humans as a solution. This work aimed to integrate what has emerged from the literature with the well-known NIST framework to provide practitioners and scholars with an additional tool to analyze their cyber resilience status.

- RQ3: Does the effectiveness of selected cybersecurity practices differ in the case of internally managed or outsourced cybersecurity processes?

The purpose is to test the effectiveness of implementing cybersecurity outsourcing strategies for small and medium-sized organizations. A literature review refines the concept of cybersecurity outsourcing. Then, through an exploration of the NIST cybersecurity framework, a group of organizational cybersecurity practices was selected, and hypotheses related to their effectiveness were formulated. The hypotheses were tested involving cybersecurity experts. The study contributed to the debate on the decision-making process for choosing to outsource cybersecurity by stressing the importance of also considering managerial aspects and variables as external leverage to improve cyber resilience.

- RQ4: Leveraging human-machine interaction for cyber resilience
    - RQ4.1: How are conceptually grounded design elements for digital intelligent agents (DIAs)?
    - RQ4.2: How to enhance cybersecurity through DIAs?

The purpose was the construction of an integrated conceptual architecture and the development of a taxonomy for the design of intelligent digital agents. The architecture and taxonomy were validated following a methodology based on literature and focused case studies analysis. Then, a technical solution was proposed, and an experiment was conducted to answer RQ4.2. A digital intelligent assistant to support the operator who is dealing with high cyber risk situations was developed and tested.

Table 1 details the research structure of this thesis. Figure 1 conceptually sketches the research activities conducted in the thesis to achieve the results described above. Finally, to understand the nature of the research and the contributions made the reader can refer to Figure 14 given at the end of the thesis.

Table 1 - Research Structure

RESEARCH STRUCTURE			
STEP 1	**RQ1: Which are the approaches to enhance cyber resilience?**		
	RESEARCH PURPOSE	**RESEARCH STRATEGIES**	
	Exploration The purpose is a cross-domain exploration of the literature on the topic of cyber resilience. The review highlights current approaches for enhancing cyber resilience. Possible research gaps are uncovered.	A Systematic Literature Review following Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guideline uncovers research gaps.	
STEP 2	**RESEARCH DIRECTION: Which are the managerial and operational opportunities to increase the positive role of humans in enhancing the cyber resilience of cyber socio-technical systems?**		
	RESEARCH QUESTIONS	**RESEARCH PURPOSE**	**RESEARCH STRATEGIES**
	A human-centric cyber resilience **RQ2.1:** Which are the human factors involved in cybersecurity? **RQ2.2:** How does each factor contribute as a weakness or opportunity to cyber resilience?	**Theory Building** The purpose is the identification of human factors as positive or negative variables in cybersecurity, identifying the linkages between these and the available references, with clear motivations for the identified relationships.	A systematic literature review maps the human factors involved in cybersecurity. A focused analysis of case studies is conducted. The analysis explores cyber attack stories reported in research articles that addressed the positive or negative contribution of human factors in cyber incidents.
	Discussing the effectiveness of organizational cybersecurity outsourcing practices **RQ3:** Does the effectiveness of selected cybersecurity practices differ in the case of internally managed or	**Theory Refinement** The purpose is to test the effectiveness of cybersecurity outsourcing strategies. A literature review refines the concept of cybersecurity outsourcing. Trough an exploration of the NIST.	A survey and a comparative quantitative study involving cybersecurity experts is conducted to test the hypothesis.

	RESEARCH QUESTIONS	RESEARCH PURPOSE	RESEARCH STRATEGIES
	outsourced cybersecurity processes?	cybersecurity framework a group of organizational cybersecurity practices is selected and hypotheses related to their effectiveness formulated. The hypotheses are tested involving cybersecurity experts	
	Leveraging human-machine interaction for cyber resilience **RQ4.1:** How are conceptually grounded design elements for digital intelligent agents (DIAs)? **RQ4.2:** How to enhance cybersecurity through DIAs?	**Theory Building** The purpose is the construction of an integrated conceptual architecture and the development of a taxonomy for the design of intelligent digital agents. The architecture and taxonomy are tested with an experiment consisting in the deployment of a DIA for cybersecurity.	Focused case studies are used to validate a conceptual architecture and a taxonomy. An experiment is conducted. A digital intelligent assistant for supporting employees during cybersecurity issue is developed and tested.

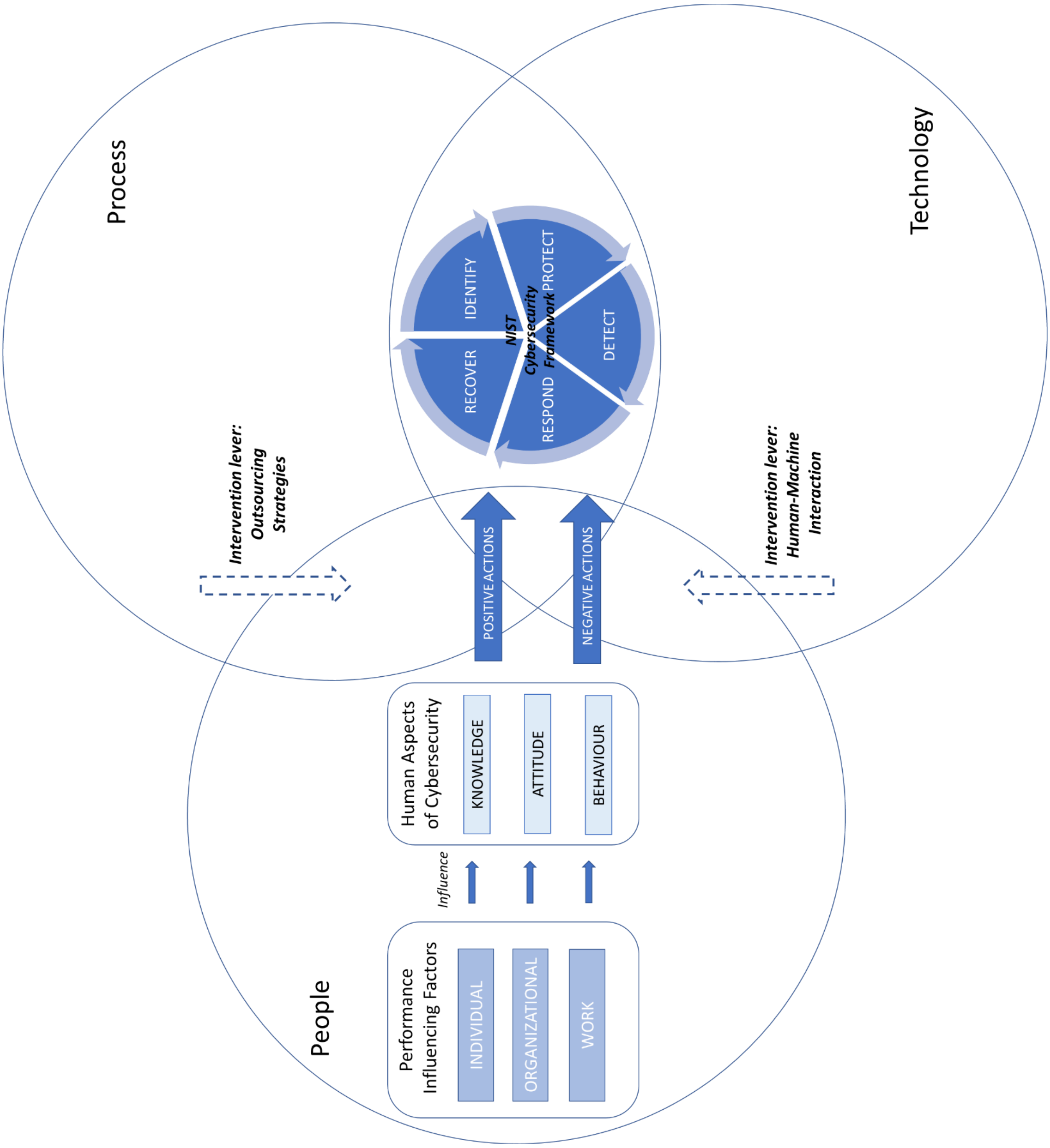

Process
Technology
People
IDENTIFY
PROTECT
DETECT
RESPOND
RECOVER
NIST Cybersecurity Framework
Intervention lever: Outsourcing Strategies
Intervention lever: Human-Machine Interaction
POSITIVE ACTIONS
NEGATIVE ACTIONS
Human Aspects of Cybersecurity
KNOWLEDGE
ATTITUDE
BEHAVIOUR
Influence
Performance Influencing Factors
INDIVIDUAL
ORGANIZATIONAL
WORK

Figure 1 - Research Framework

1.1. Thesis Outline

In the remainder of Chapter 1, the research structure of the thesis is introduced, and the research framework explained the research activities conducted. Chapter 2 describes the important concepts and definitions related to the broad research topic of cyber resilience for clarifying the context and thesis boundaries. Chapter 3 conducts a systematic literature review on cyber resilience answering RQ1. This review uncovered unexplored research streams and identified the research gaps addressed in the additional research questions detailed in the second part of the thesis. Chapter 4 answers RQ2 by detailing the role of human factors in cybersecurity and proposing how to leverage human factors to assess, manage and improve cyber resilience. Chapter 5 and Chapter 6 answer respectively RQ3 and RQ4 presenting two specific opportunities linked to human factors and cyber resilience. Chapter 5 investigates the effectiveness of cybersecurity outsourcing practices. Chapter 6 presents the digital intelligent assistant architecture, taxonomy, and a digital assistant prototype for supporting workers in case of a cyber attack. Chapter 7 discusses the obtained results and summarizes the conclusions of the thesis.

2. Conceptual background

Today hyper-connected environment has led on one side to an appreciable increase in productivity, efficiency, and system integration, on the other side it has increased the number of potential risks. Organizations are now highly dependent on data and information from their integrated systems [2]. This dependence has highlighted how a cyber threat can be substantial in terms of continuity of business operations, theft of confidential information, and reputational harm. For these reasons, cyber resilience has become a top priority for organizations that find themselves working in an increasingly dynamic and real-time optimized network [3]. Over the past decade, attacks have not only grown in numbers but also in power and sophistication. Now, malicious actors can not only steal our information but also potentially cause physical harm. The case of the well-known Stuxnet attack, first uncovered in 2010, is exemplified. Stuxnet is a computer virus probably created and spread by the U.S. government as part of Operation Olympic Games, which consisted of a series of "digital attacks" against Iran in cooperation with the Israeli government [4], [5]. The purpose of the software was to sabotage Iran's Natanz nuclear power plant. Specifically, the virus was meant to disable the plant's centrifuges, hindering the detection of malfunctions and the virus' detectability [6]. Stuxnet targeted PLCs, software-programmable hardware components critical to the automation of the plant's facilities, particularly those used to control the centrifuges (used to separate nuclear materials such as enriched uranium). The feature that impressed experts was the level of sophistication of this software, which showed that those who designed the program were familiar with the computer network environment in use at the power plants. The Stuxnet "worm" has the potential to herald an unsettling new era for cybersecurity. It was the first cyberweapon to illustrate exactly how vulnerable the digital world is. The Stuxnet "worm" has the potential to usher in a disturbing new era for cybersecurity. What emerges and will be detailed in later chapters of this thesis, is the now imperative integration between the concepts of systems security and the safety of the individuals involved [7].

Moreover, the rapid and worldwide expansion of the coronavirus pandemic highlighted the vulnerability of traditional cybersecurity systems [8]. IT has played a positive significant role in all activities, serving as the focal point of operations in healthcare, business, education, industry, and more [9]. However, there are several drawbacks, such as increased cybersecurity threats and hazards, performance problems because of a massive increase in workload, and business continuity. By increasing remote work, IT networks and systems have become more vulnerable to threats causing damage to business operations, inflicting substantial costs, and compromising the reputation of companies. As a result, the exponential increase in the use of smart working, using personal devices, home networks, and collaboration platforms, has contributed to expanding, from a security perspective, the attack surface exploitable by malicious actors [10]. Companies are reviewing their IT spending, and planning a significant increase in investment in IT security [11], [12]. So, abetted by the health emergency, we can conclude that the increased use of IT platforms and technologies such as cloud computing, data analytics, and virtualization has led to IT security being considered a top investment priority in organizations [13]. To defend against this new vulnerability, organizations have adopted techniques to combat cybersecurity breaches, including firewalls, encryption techniques, access control mechanisms, intrusion detection systems, and continuous workforce training.

2.1. Cybersecurity definition(s)

As the internet has grown, even faster have grown cybersecurity issues, data privacy, and online rights. Governments, organizations, and researchers have been playing catch-up, defining cybersecurity in different ways, and taking different approaches to regulating and enforcing the rules of cyberspace, without reaching an international consensus.

On this perspective, experts from the European Union Agency for Network and Information Security (ENISA) produced an interesting report in 2015 [14] emphasizing the importance of cybersecurity standardization to raise the level of cybersecurity. In particular, the report proposes a revision of the definitions of the term "Cybersecurity" (or "cybersecurity").

What emerges is the difficulty of providing a definition that encompasses everything cybersecurity covers. Accordingly, a contextual definition that is relevant, appropriate, and already used by academics and organizations is explored in this thesis.

Cybersecurity refers to the protection of cyberspace, which is the collection of connections and relationships between objects that can be accessed through a wide-area network of telecommunications, and the set of objects that have interfaces that enable remote control, remote access to data, or participation in control operations within that cyberspace.

First, researchers now agree that with the term cybersecurity we include multiple domains. Specifically, we find within the definition the domains: communications security, operations security, information security, military security, and physical security. For a more extensive definition of each of these concepts, we refer to the glossary in the appendix. Therefore, the definition proposed by, for instance, The Oxford English Dictionary which defines "Cybersecurity" as "The state of being protected against the criminal or unauthorized use of electronic data, or the measures taken to achieve this" is considered outdated. Indeed, this statement does not include operational errors, human errors, or manipulation of physical assets, such as industrial settings, critical infrastructures, etc.

According to the report, each definition of Cybersecurity contains a series of components. They represent all these characteristics in a diagram shown in Figure 2. The definition used in this thesis integrates the NIST National Institute of Standards and Technology [15] definition of cybersecurity with ISO/IEC JTC1/SC27 27000 [16] definition of information security. We have shaded in color the components included in this thesis definition (Figure 2). The definition which the author will refer to is "Cybersecurity ensures confidentiality, availability, and integrity and is achieved through the application and management of a set of controls, including policies, processes, procedures, organizational structures, software, and hardware to prevent damage to, protection of, and restoration of cyberspace and the information contained therein.”

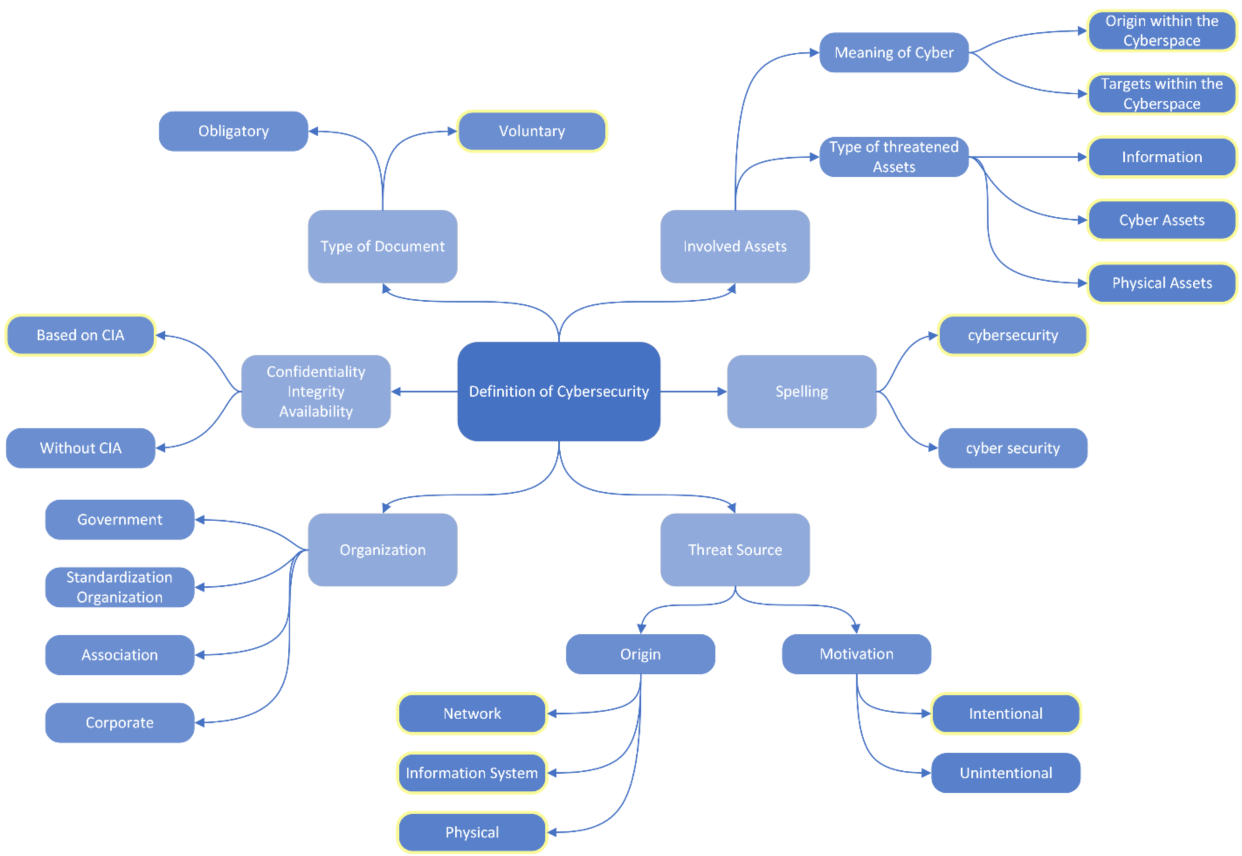


Figure 2 – Definition of Cybersecurity Diagram according to ENISA [14]. *The components included in this thesis definition are outlined in yellow.*

Finally, terms such as “Confidentiality, Integrity, and Availability” used in the definition refer to the CIA paradigm. The description of each element is given below [13], [17], [18]:

- *Confidentiality*: users' sensitive and personal data mustn't be disclosed to unauthorized persons, therefore, it is necessary to apply specific levels of access for those who are authorized to access it.
- *Integrity*: aims to protect data in the system, therefore, information from various organizations and sources must not be changed or altered by an unauthorized entity under any circumstances.
- *Availability*: refers to the actual availability of data; If a customer wishes to access their account or log in to a system, they should be allowed to access or log in at any time.

2.2. Cyberattacks

A cyber attack means any maneuver employed by individuals or organizations that affects computer systems, infrastructure, networks, and/or electronic devices through malicious acts aimed at the theft, alteration, or destruction of specific targets by breaching susceptible systems [16].

As we can see, the definition of a cyber attack is broad and generic. It refers to a set of activities, which can be applied to the entire cyberspace. The agent performing a cyberattack is called an attacker. In contraposition to this term, the agent in charge of preserving the target system is called the defender.

The attacker has a deep knowledge of the computer and the network and intervenes on memories to which he does not have legal access to steal or alter data. Specifically, it is the one who commits a crime in the context of cyber crime. Cyber crime is defined as any computer crime involving a device, computer, network, etc. An attacker, whether a freelancer, a government employee, or part of a military corps, is able to recognize vulnerable computer systems that lack appropriate security measures. Once a vulnerability is discovered, the attacker can infect the system with malicious code and gain remote control of it and then retrieve its contents or use it to damage other computers. A vulnerability is a weakness in an IT system that can be exploited by an attacker to deliver a successful attack. They can be brought about by flaws, features, or user mistakes, and attackers will try to take advantage of any of them, frequently combining one or more, to accomplish their objective [19]. It is worth mentioning that for the purpose of this thesis it has been considered unnecessary a distinction between attackers with “positive connotations” (e.g., hacktivist) and attackers with “negative connotations” (e.g., malicious hacker, espionage hacker). On the other side, the defender must deal with attacks in a context characterized by vulnerabilities. The defender is aware of known vulnerabilities for which a defender strategy is developed. However, there exist also unknown vulnerabilities which are the ones that have not been exploited yet by the defender. According to the data available in the Common Vulnerabilities and Exposures (CVE) database, only in 2021, there have been reported around fifty-five vulnerabilities per day. CVE’s goal is to identify, define and catalog publicly disclosed vulnerabilities related to security issues. It enables businesses to analyze and prioritize vulnerabilities, compare their severity, and monitor their cybersecurity posture over time. Figure 3 shows the number of vulnerabilities reported to CVE to date regardless of whether they were exploited or not.

For cyber attacks, too, we can make a parallel observation. An attack is referred to as known when it has been identified, analyzed, and the vulnerability used detected. The following are the statistics that the Italian Association for Information Security (CLUSIT) details each year in a report on the state of cybersecurity in Italy [20] The latest report from October 2022, shows how attacks around the world have increased by 8.4% percent over the previous year and are getting more serious. Attacks are growing in quantity and sophistication, reporting an average of 190 attacks per month. Attacks classified by CLUSIT researchers occurred mainly in the American continent. However, attacks in Europe have grown, exceeding one-fifth of the total (26%, up from 16% in the previous year). In 2022, 78% of detected attacks experienced a "high" or “critical” impact, up from 50% two years ago. Finally, cybercrime is confirmed as the main motivation leading to attacks, with an increasing number of Cyber Espionage and Information Warfare attacks. From the perspective of attack targets, the most affected sectors have been Government Infrastructure, Healthcare, ICT, Telecommunication, Multiple Small Targets and Energy.

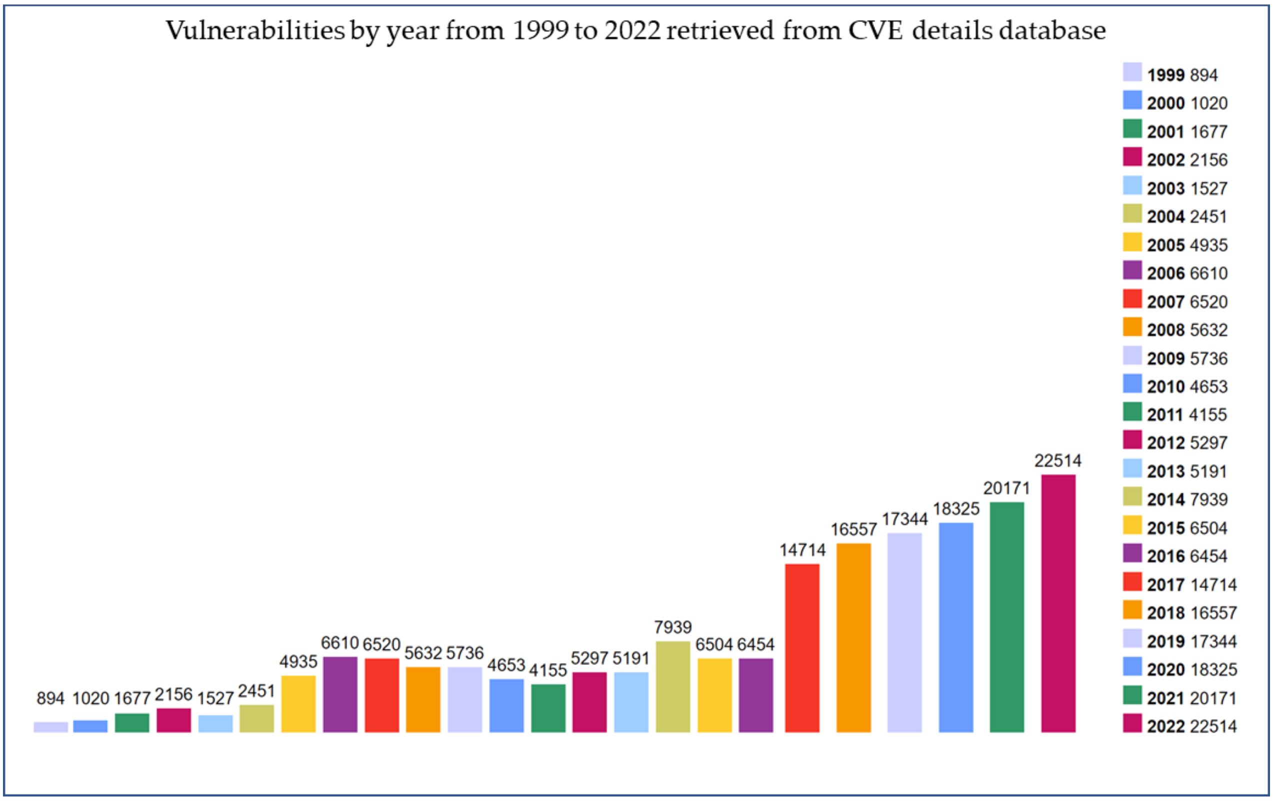


Figure 3 – CVE vulnerabilities from 1999 to 2022 retrieved from [21]

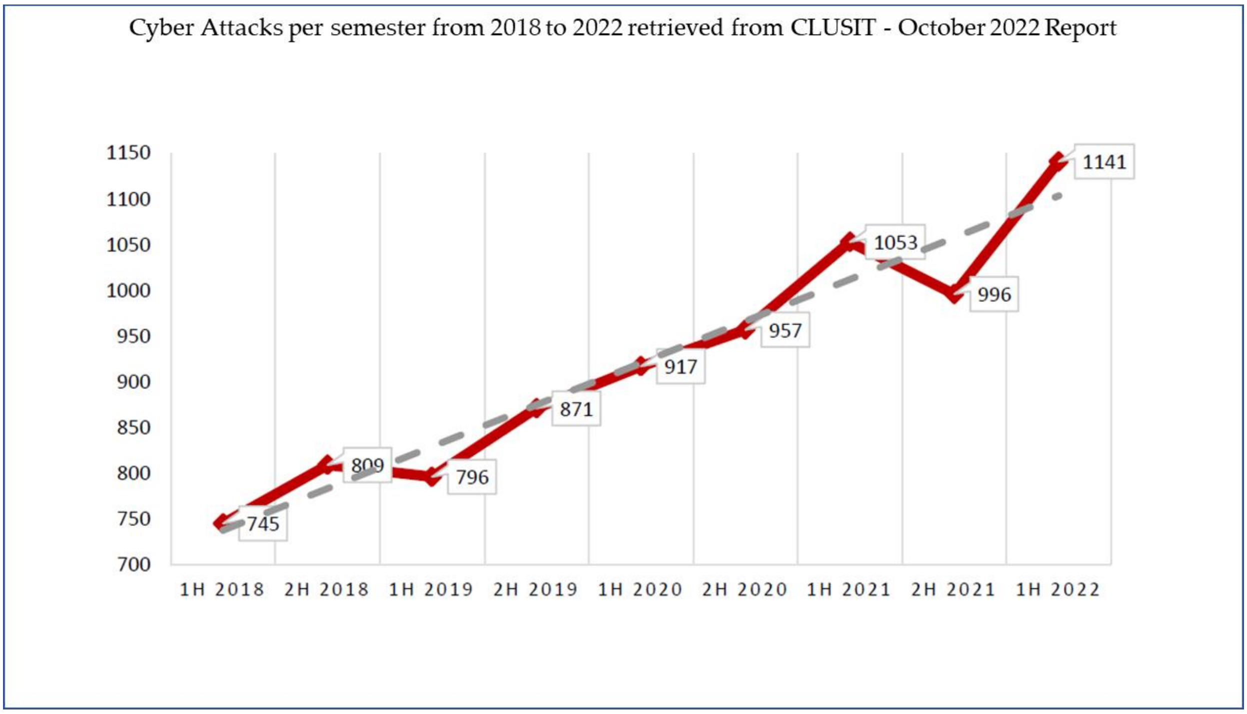


Figure 4 – Cyber attacks registered globally from January 2018 to June 2022 retrieved from [20]

Finally, turning to look at the attack techniques used to conduct a cyber attack, each has different specific characteristics, targets, and levels of sophistication. Table 2 shows an overview of common cyber attack techniques. For each of them, a revised definition is suggested considering those proposed in the literature.

Table 2 - Main cyberattacks techniques

Cyber attack	Description	Reference
Denial-of-service (DoS) attack and Distributed-denial-of-service (DDoS) attack	DoS attacks aim at deteriorating the communication channels to prevent information exchange, usually either sensor data or control commands, between components of the system. In this type of attack, a huge volume of data is transmitted to the network to make the server busy, thereby disrupting normal services.	[22]–[24]
Jamming attack	It is considered a kind of DoS attack and the most basic form of radio attack on a wireless communication system. It aims to disrupt a CPS preventing the control system from accessing current sensor measurements.	[25]–[27]
False data injection (FDI) attack	FDI attacks are considered among integrity attacks. In an FDI attack, an adversary could access and modify the physical system's state, sensor data, or control commands by introducing arbitrary errors and fake information. FDI attacks have been widely used against grid state estimation and energy management systems to disrupt the energy distribution.	[26], [28], [29]
Man-in-the-middle (MITM) attack	MITM attacks are part of the integrity attacks. Specifically, a MITM attack is an unauthorized interception and potential alteration of communication between two parties by an unauthorized third party. A representative attack of this kind in	[30]–[32]

Cyber attack	Description	Reference
	CPSs might disrupt data integrity establishing communication between the physical plant to the feedback controller through the network. Then, the original data is modified by the adversary sending false data to the feedback controller which processes it accordingly.	
Stealthy attack	Stealthy attacks are sophisticated and potentially dangerous attacks. Usually conducted by skilled attackers who can penetrate control networks and then manipulate sensor readings or control signals persistently until the system crashes, while still keeping themselves undetected by following the expected behavior of the system closely.	[30], [33], [34]
Malware attacks: ransomware, viruses, spyware, worm	A malware attack is a common cyberattack where malicious software can perform a variety of tasks and damage CPSs in multiple ways. It encompasses specific types of attacks such as ransomware, viruses, and worm. Ransomware denies or restricts access to victim files and demands a payment from the victim to restore access. Viruses are a type of malware that propagates by inserting a copy of itself and becoming part of another program. They spread from one computer to another, leaving infections as they travel. They can range in severity from causing mildly annoying effects to damaging data or software and causing denial-of-service (DoS) conditions. Unlike Viruses, Worms are spread via software vulnerabilities or phishing attacks. Worms and Viruses can modify and delete files, inject malicious software, and steal data. Finally, spyware is a program that accesses a user's personal information and then transmits the information to an adversary for misuse.	[35]–[37]

Cyber attack	Description	Reference
Phishing – Smishing attack	A type of social engineering attack in which victims are targeted through a link, usually found in an e-mail a text message sent to them. The link, once clicked, may contain a malware, or may trigger further communications requesting personal information from the user. A "spear-phishing" attack is a more sophisticated type of phishing attack that targets a specific individual, typically using publicly available information about that individual, thereby personalizing the communication in an attempt to increase the likelihood that the targeted person will click on the link.	[38], [39]
Pretexting	A type of social engineering attack that involves a situation, or pretext, created by an attacker in order to lure a victim into a vulnerable situation and to trick them into giving private information, specifically information that the victim would typically not give outside the context of the pretext	[40]

2.3. Cyber Resilience

The idea of resilience was explicitly presented, referring to ecological problems as the persistence of interactions within a system, and it was assessed by the system's capacity to absorb change-state variables, driving variables, and parameters while being persistent [41]. [42] identifies four primary disciplinary viewpoints that have consistently addressed resilience issues. The suggested application domains for resilience include societal, organizational, economic, and engineering. Resilience is always viewed as an intrinsic capacity of individuals, communities, and environments in the social domains. The issue of resilience originated in the organizational domain because of organizations' need to adjust to a quickly changing business environment. Organizational resilience is described as the capacity to maintain a steady state following a disruptive event or as the capacity and speed of an organization to resume regular operations following a disruptive event [43]. Instead, [44] provides a fairly precise definition of economic resilience as the ability to reconfigure and alter the structure of an economy (firms, industries, technologies, institutions), in order to sustain an acceptable growth path in production, employment, and wealth through time. Lastly, some engineering domain definitions were proposed by [45]. In [45] the authors state that *"A system is resilient if it can adjust its functioning prior to, during, or following events* *(changes, disturbances, and opportunities), and thereby sustain required operations under both expected and unexpected conditions".* According to [46], a resilient system must be able to:

- *Respond*: the ability to know what to do and be capable of responding to regular and irregular changes and opportunities by activating actions or adjusting current operations.
- *Monitor*: the ability to know what to look for (e.g., indicators) and being able to monitor a system's performance and environment.
- *Learn*: the ability to know what has happened and be able to learn from experience.
- *Anticipate*: the ability to know what to expect and being able to anticipate disruptions, opportunities, novel demands, or constraints by observing how factors interact and influence each other.

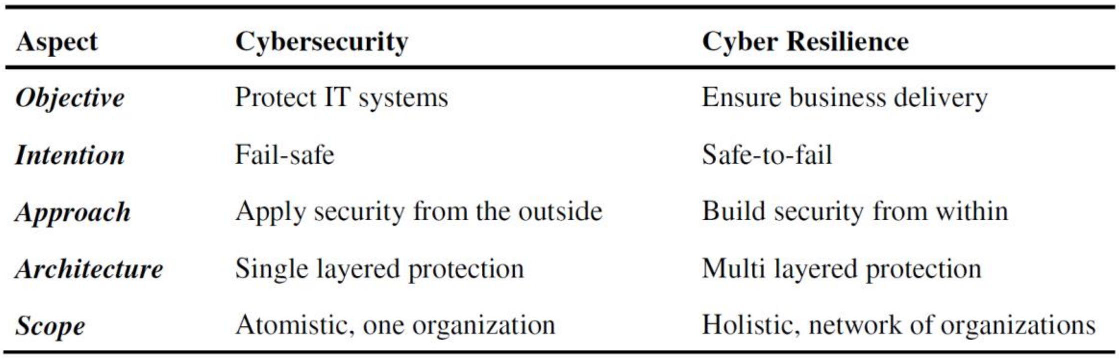
For the purpose of this thesis, the author will refer to the last mentioned definition of resilience engineering encompassing the idea proposed by [47] as a progressive shift from socio-technical systems to cyber socio-technical systems.

Today increasing digitalization and autonomation of work processes requires to include interconnected cyber-technical artifacts in socio-technical systems.

In this scenario, a cybersecurity issue does not only refer to data or information hacking, but it can lead to a modification of physical world processes, involving the entire system and causing tangible damages. Even though systems are becoming increasingly software-centric, intrusions may still have physical implications. In this scenario, evaluating resilience as it is commonly described makes it impossible to investigate the safety and system dependability. It is vital to include not only failures strictly connected to system physical components but also anomalies in the IT environment. As a result, a more precise notion of cyber resilience began to be considered, experimenting an increase in its usage in the last few years, giving tools for assessing and quantifying the resilience of cyber socio-technical systems. Specifically, [1] offers one of the most cited definitions of cyber resilience, stating that cyber resilience is *"the ability to continuously deliver the intended outcome despite adverse cyber events caused by humans and nature"*. The term ability refers to the ability to continuously deliver the intended outcome independently if it is a nation, an organization, or a specific IT system. The concept of continuously implies that the capacity to deliver the intended output should be able to operate even when regular delivery mechanisms fail, during a crisis, or following a security breach. It is an ongoing practice, not a one-time effort. It is the ability to continually adapt to new or potential risks. The intended outcome refers to what the system is intended to achieve, such as the goals of a business or the services delivered by an online service. Finally, the term adverse events refers according to Bjork to all events, unintended and intended threats, that impact the confidentiality, integrity and availability of IT systems. Other definitions exclude those not deliberate cyber events from the definition by stating that *"cyber resiliency means a system's ability to tolerate a cyber attack"* [48]. Other definitions, however, are more specific from the perspective of the infrastructure or sector considered. This is the case of [49], which defines cyber resilience as *"the capacity of a power enterprise to maintain its core purpose and integrity in the face of Cyberattacks."* Similar are the works of [50] or [51] who present works on cyber resilience in critical energy infrastructure proposing comparable definitions. Finally, many studies that deal with the topic investigate it from a perspective focused on measuring cyber resilience. There are numerous works focused on risk assessments, emergency preparedness, cyber resilience metrics or models, many of which are collected in the review conducted during this thesis track and presented in the following chapter.

2.3.1. Cybersecurity and Cyber resilience

The major difference between cybersecurity and cyber resilience lies in the objective and scope. The objective of cybersecurity is to keep IT systems alive while cyber resilience during adverse events focuses on also keeping business goals intact. Cyber resilience requires a holistic approach that includes information, technologies, people, and processes. On the other hand, cybersecurity focuses on a single unit of analysis within an environment.
However, we can conclude that the cybersecurity approach includes within it multiple components and characteristics of cyber resilience. The following is a summarized description of the differences between cybersecurity and cyber resilience proposed by [1].



Aspect	Cybersecurity	Cyber Resilience
Objective	Protect IT systems	Ensure business delivery
Intention	Fail-safe	Safe-to-fail
Approach	Apply security from the outside	Build security from within
Architecture	Single layered protection	Multi layered protection
Scope	Atomistic, one organization	Holistic, network of organizations

Figure 5 - Cybersecurity vs Cyber Resilience retrieved from [1]

2.4. Cybersecurity Frameworks and Standards

As mentioned in the previous paragraph, cyber resilience is an organization's ability to deliver intended outcomes even under adverse cyber circumstances. However, attaining cyber resilience requires organizations to gather and analyze high-quality intelligence about both the cyber landscape and organizational circumstances. As a result, a multitude of legislative acts, standards, and framework has been introduced, attempting to combat both cybercrime and privacy infringements and thereby attain cyber resilience [52]. A cybersecurity standard is a set of guidelines or best practices that businesses may apply to strengthen their cybersecurity posture. Cybersecurity standards may assist organizations in identifying and implementing suitable steps to safeguard their systems and data from cyber adverse events. Standards can also assist users to respond to and recover from cybersecurity problems. Instead, a cybersecurity framework is a set of rules, guidelines, and best practices for managing risks in the digital world. It is generally applicable to all organizations, regardless of their size, industry, or sector.

Below are the main standards, guidelines, and frameworks used to achieve the objectives of this thesis.

2.4.1. Standards and Regulations

- ISO/IEC 27000 is a standard family comprising sixty standards covering a broad spectrum of information security issues. The standard provides an overview of information security management systems (ISMS) providing terms and definitions commonly used [16].

- ISO/IEC 27001 is an international standard for information security that provides a framework for managing sensitive company information. The Standard includes requirements for developing an ISMS (information security management system), implementing security controls, and conducting risk assessments [53].

- ISO/IEC 27002 is the code of practice for information security management. It provides guidance and recommendations on how to implement security controls within an organization [54].
- ISO/IEC 27031 describes guidance on IT disaster recovery programs and related activities. It describes provides a framework of methods and processes to identify and specify all aspects or improving an organization's ICT readiness to ensure business continuity [55].

- ISO/IEC 27032 is an internationally recognized standard that provides guidance on cybersecurity for organizations. The Standard is designed to help organizations protect themselves against cyber attacks and manage the risks associated with the use of technology. It is based on a risk management approach and guides how to identify, assess, and manage cyber risks. The Standard also includes guidance on incident response and recovery [56].

- ISO/IEC 27701 specifies the requirements for a PIMS (privacy information management system) based on the requirements of ISO 27001. It is extended by a set of privacy-specific requirements, control objectives, and controls [57].

- GDPR (General Data Protection Regulation) is a regulation in EU law on data protection and privacy in the European Union. The European community's goal was to strengthen the protection of the personal data of European Union (EU) citizens, both within and outside the EU borders, by giving citizens back control of their personal data, unifying and homogenizing privacy regulations within the EU.
  The relationship between GDPR and cybersecurity lies in the introduction of a new methodological approach called risk-based. This approach is based on the implementation of data protection and data security measures to protect data from attacks that can alternate its CIA, confidentiality, integrity, and availability.

This type of security must be planned by the design of data processing *(data protection by design*). However, the purposes of data and data security cannot exceed privacy rights and thus cannot exceed keeping others' confidential information under control *(data protection by default)* [58].

- NIS (Network and Information Security) directive is a European Union directive that contains a series of legislative measures aimed at creating a common level of network security and information systems in general within the European Union. The main purpose of this European legislation is to ensure that each EU member state improves its ability to manage network security; that all states can recognize and manage the most serious risks and errors of operators and providers of digital services. In May 2018, Italy also adopted the NIS directive. Each state is required to adopt a national-level cybersecurity strategy. The European Union oversees adopting all the necessary tools for a common and cooperative network, which will be fast and effective, so that trust will develop between the states. The latter will also be tasked with appointing competent national authorities, as well as figures who will be given responsibility for monitoring incidents in this area: (e.g., a person responsible for the Computer Security Incident Response Team - CSIRT) [59].

2.4.2. Frameworks

- NIST Cybersecurity Framework

The National Institute of Standards and Technology (NIST) is a technology management organization of the U.S. government. NIST provided a "Framework for Improving Critical Infrastructure Cybersecurity" [60]. This framework is among the most widely used worldwide and provides an approach to securing a computer network of any size. The framework provides a set of standards, guidelines, and best practices for managing cybersecurity risks.
The framework consists of the Core framework, Implementation Tiers, and Profiles.

The Core framework encapsulates the set of activities and deliverables useful to the cybersecurity management process. The goal of this component is to foster communication and collaboration among multidisciplinary teams by avoiding technical language in favor of a simpler, more intuitive language. The Core consists of three parts: Functions, Categories, and Subcategories (see Figure 6).
The Core Framework also provides references that connect each subcategory to safety practices that are known to be necessary for compliance with industry standards (ISO, SP800-53r4, COBIT-5, SANS20, and others) or with current general laws (EU Regulation 2016/679 General Data Protection Regulation, EU Directive 2016/1148 NIS).
There are five functions in the core named as follows: Identify, Protect, Detect, Respond, e Recover.

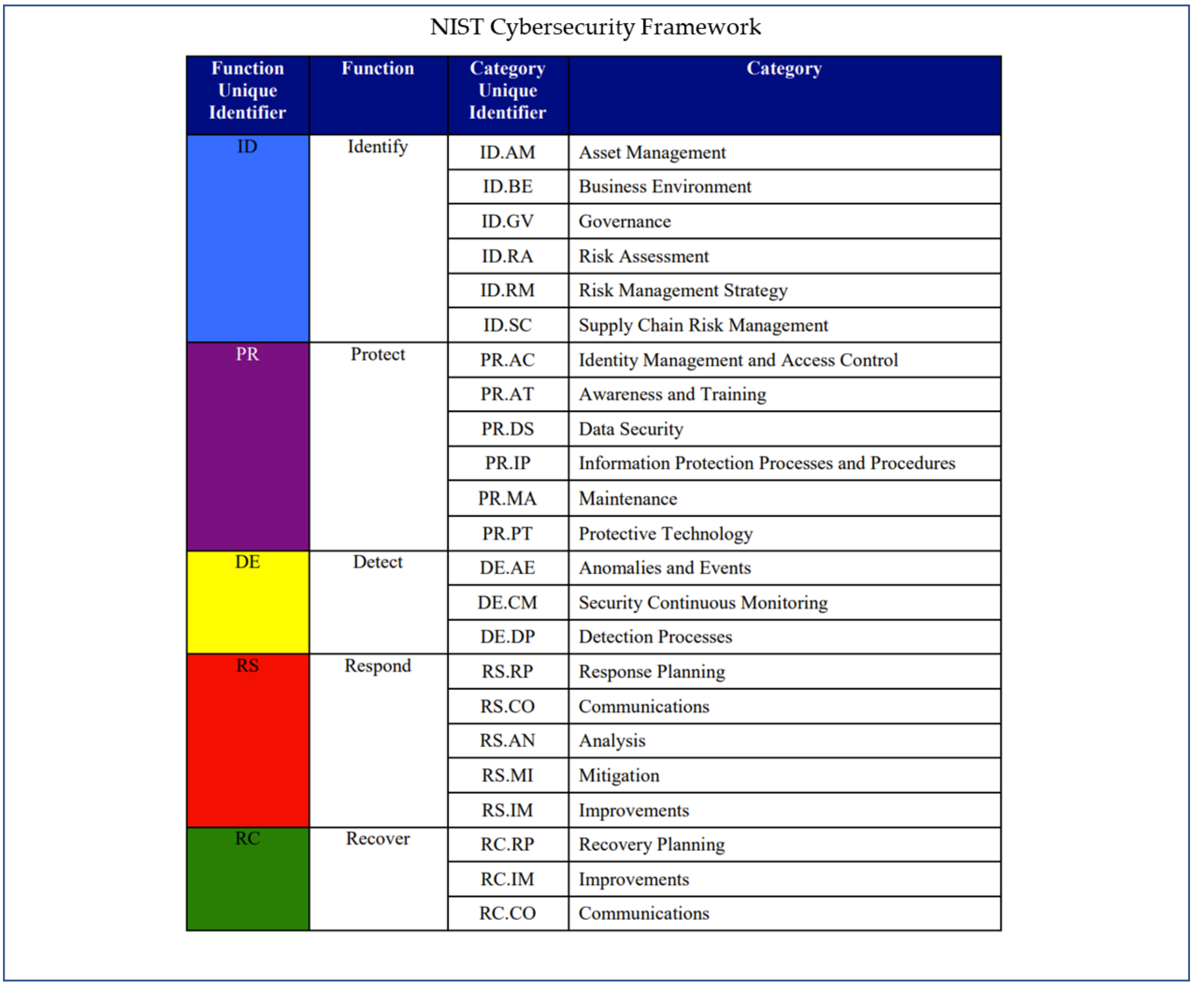
- *IDENTIFY*: Develop an organizational understanding to manage cybersecurity risk to systems, people, assets, data, and capabilities.

- *PROTECT*: Develop and implement appropriate safeguards to ensure the delivery of critical services.
- *DETECT*: Develop and implement appropriate activities to identify the occurrence of a cybersecurity event.
- *RESPOND*: Develop and implement appropriate activities to act regarding a detected cybersecurity incident.
- *RECOVER*: Develop and implement appropriate activities to maintain plans for resilience and to restore any capabilities or services that were impaired due to a cybersecurity incident.

For each of these, there are several categories assigned for a total of twenty-three.

Categories are used to report on cybersecurity goals an organization should pursue. To this end, these are designed without being overly detailed to provide a broad scope. Indeed, topics such as physical security, logical security, personal data security, or business outcomes are addressed. Finally, in turn, the categories have subcategories. Subcategories represent the deepest level of abstraction in the Core. There are 108 subcategories, which are outcome-driven statements that provide considerations for creating or improving a cybersecurity program.

Outcome-driven means that guidance is given by results. However, this approach does not dictate how an organization should achieve those outcomes. Each implements the solution based on its own risk and needs as an organization.



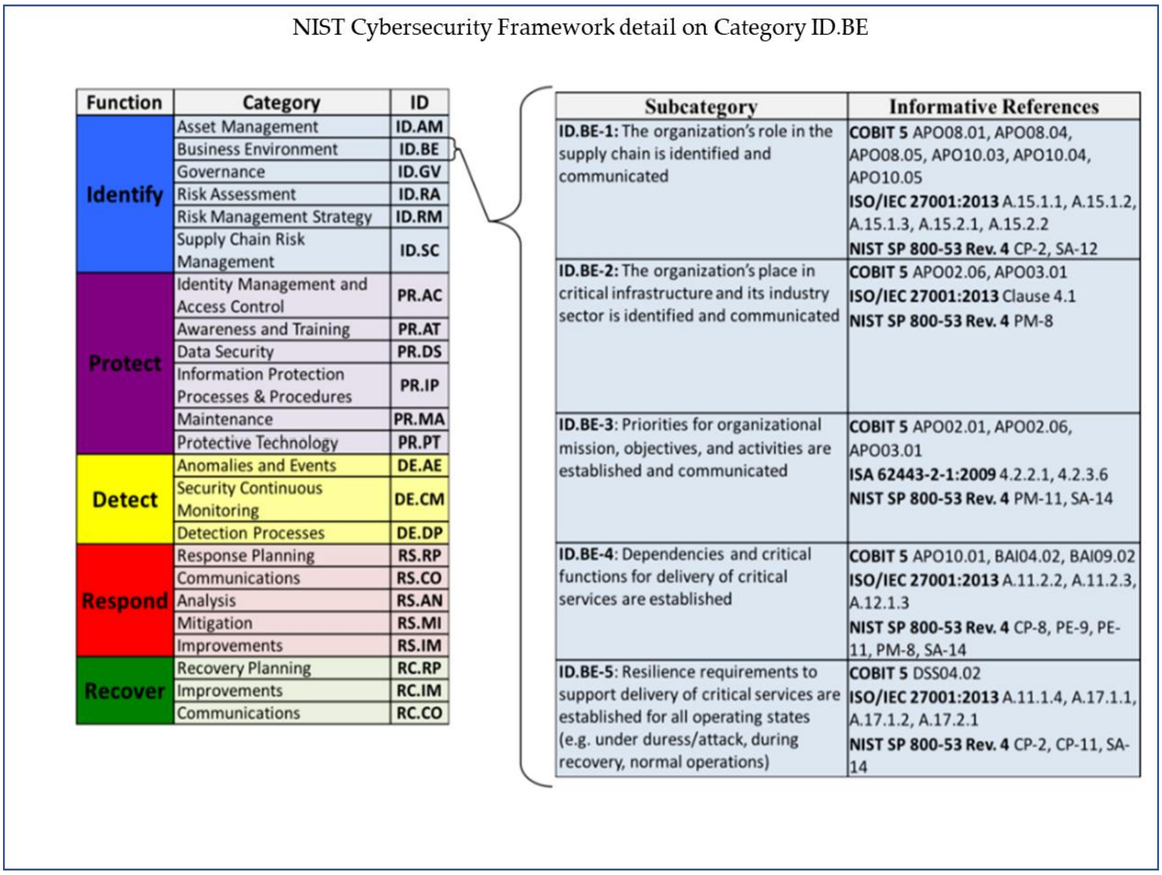
NIST Cybersecurity Framework

Function Unique Identifier	Function	Category Unique Identifier	Category
ID	Identify	ID.AM	Asset Management
		ID.BE	Business Environment
		ID.GV	Governance
		ID.RA	Risk Assessment
		ID.RM	Risk Management Strategy
		ID.SC	Supply Chain Risk Management
PR	Protect	PR.AC	Identity Management and Access Control
		PR.AT	Awareness and Training
		PR.DS	Data Security
		PR.IP	Information Protection Processes and Procedures
		PR.MA	Maintenance
		PR.PT	Protective Technology
DE	Detect	DE.AE	Anomalies and Events
		DE.CM	Security Continuous Monitoring
		DE.DP	Detection Processes
RS	Respond	RS.RP	Response Planning
		RS.CO	Communications
		RS.AN	Analysis
		RS.MI	Mitigation
		RS.IM	Improvements
RC	Recover	RC.RP	Recovery Planning
		RC.IM	Improvements
		RC.CO	Communications

Figure 6 - NIST Cybersecurity Framework retrieved from [60]

Let us consider the *"Identify"* function (see Figure 7). For the above, the various categories include *"Business Environment"* which, in turn, has several subcategories. As an example, one of them is, *"Priorities for organizational mission, objectives, and activities are established and communicated."*

As can be seen then, the subcategories are actual "statements" of the organization's achievements. Finally, each subcategory is equipped with a complete set of references to additional technical documents, called Informative References, which can be used, as reference elements to be followed to achieve the objectives set in the subcategories.



NIST Cybersecurity Framework detail on Category ID.BE

Function	Category	ID
Identify	Asset Management	ID.AM
	Business Environment	ID.BE
	Governance	ID.GV
	Risk Assessment	ID.RA
	Risk Management Strategy	ID.RM
	Supply Chain Risk Management	ID.SC
Protect	Identity Management and Access Control	PR.AC
	Awareness and Training	PR.AT
	Data Security	PR.DS
	Information Protection Processes & Procedures	PR.IP
	Maintenance	PR.MA
	Protective Technology	PR.PT
Detect	Anomalies and Events	DE.AE
	Security Continuous Monitoring	DE.CM
	Detection Processes	DE.DP
Respond	Response Planning	RS.RP
	Communications	RS.CO
	Analysis	RS.AN
	Mitigation	RS.MI
	Improvements	RS.IM
Recover	Recovery Planning	RC.RP
	Improvements	RC.IM
	Communications	RC.CO

Subcategory	Informative References
ID.BE-1: The organization's role in the supply chain is identified and communicated	**COBIT 5** APO08.01, APO08.04, APO08.05, APO10.03, APO10.04, APO10.05 **ISO/IEC 27001:2013** A.15.1.1, A.15.1.2, A.15.1.3, A.15.2.1, A.15.2.2 **NIST SP 800-53 Rev. 4** CP-2, SA-12
ID.BE-2: The organization's place in critical infrastructure and its industry sector is identified and communicated	**COBIT 5** APO02.06, APO03.01 **ISO/IEC 27001:2013** Clause 4.1 **NIST SP 800-53 Rev. 4** PM-8
ID.BE-3: Priorities for organizational mission, objectives, and activities are established and communicated	**COBIT 5** APO02.01, APO02.06, APO03.01 **ISA 62443-2-1:2009** 4.2.2.1, 4.2.3.6 **NIST SP 800-53 Rev. 4** PM-11, SA-14
ID.BE-4: Dependencies and critical functions for delivery of critical services are established	**COBIT 5** APO10.01, BAI04.02, BAI09.02 **ISO/IEC 27001:2013** A.11.2.2, A.11.2.3, A.12.1.3 **NIST SP 800-53 Rev. 4** CP-8, PE-9, PE-11, PM-8, SA-14
ID.BE-5: Resilience requirements to support delivery of critical services are established for all operating states (e.g. under duress/attack, during recovery, normal operations)	**COBIT 5** DSS04.02 **ISO/IEC 27001:2013** A.11.1.4, A.17.1.1, A.17.1.2, A.17.2.1 **NIST SP 800-53 Rev. 4** CP-2, CP-11, SA-14

Figure 7 - NIST Cybersecurity Framework Detail retrieved from [61]

The second item that composes the framework are the Profiles. Profiles encapsulate within them:

- The organizational requirements and objectives
- The organization's risk appetite.
- The resources used relative to the Core's desired outcomes.

Profiles can be used as an opportunity to improve security status by comparing a current profile (current profile) with the desired profile (target profile). To develop a profile, an organization must examine each of the sub-categories and, based on its objectives and risk assessment, determine which ones are applicable in its context. The sub-categories may be integrated with additional practices not envisaged by the Framework to fully manage the risk. The current profile can then be used to prioritize and measure progress toward the desired profile. Profiles can also be used to conduct a self-assessment or to communicate one's level of cyber risk management within or outside the organization.

The third item are the Implementation Tiers. They function as benchmarks as to how well organizations are following the rules and recommendations of the Cyber-Security Framework. The implementation tiers are as follows, with one being the lowest and four being the highest:

1. *Partial*: If a company's cybersecurity risk management model does not systematically account for cyber risk or environmental hazards. Ad hoc procedures are frequently used to manage cyber risk. At the organizational level, there is a low level of risk awareness. There are no procedures in place for sharing cybersecurity-related information with outside organizations.
2. *Informed*: If a company has internal processes that consider cyber risk but do not apply to the entire organization. Although there is a reasonable amount of knowledge of cyber risk, there are not any widespread management methods that incorporate all organizational levels. Although the business is aware of its place in the reference environment, there is little active information exchange when it comes to cybersecurity occurrences.
3. *Repeatable*: If an organization's cyber risk management model is clearly defined, approved, and frequently updated depending on the results of the risk management process, then it is repeatable. All organizational levels practice cyber risk management, and staff members receive training to manage the responsibilities placed on them. With other entities active in the same ecosystem, the company frequently trades cybersecurity-related information.
4. *Adaptive*: The cyber risk management model of an enterprise is adaptable if it periodically modifies its cybersecurity protocols using experience and risk indicators. Through an adaptive process, the company can respond to complex attacks and continuously adapt to emerging threats. Information is continuously and instantly exchanged with other actors taking part in the same ecosystem.

- Italian National Framework for Cybersecurity and Data Protection

In 2015 the Italian cybersecurity national lab (CINI) along with the Cyber Intelligence and Information Security Center of Sapienza University of Rome (CIS) developed the National Framework for Cybersecurity and Data Protection [62] drawing inspiration from the NIST framework. Indeed, the Framework developed by NIST evaluates well the security of information and systems, maintaining that appropriate level of abstraction that can guarantee companies' autonomy in the 'application and contextualization of controls. However, the NIST framework is defined for critical infrastructure, introducing a level of complexity that is not suitable for most of the companies that make up the Italian enterprise ecosystem.

The National Framework expands this structure by including two new concepts: priority levels and maturity levels. These two concepts make it possible to consider the economic structure of our country consisting of a few large enterprises and critical infrastructures and multiple small and medium-sized enterprises. Thus, it makes the framework suitable for SMEs while retaining its initial focus on Large Enterprises and Critical Infrastructure.

Specifically, priority levels make it possible to support organizations in the preliminary identification of Subcategories to be implemented. This allows them to reduce their risk levels while balancing the effort they need to put into their implementation. Maturity levels, on the other hand, make it possible to provide a measure of the maturity of a security process, the maturity of implementation of a specific technology, or a measure of the number of resources expended in implementing a Subcategory. Maturity levels provide a benchmark against which each organization can evaluate its implementation of Subcategories and set goals and priorities for their improvement.

CIS and CINI have also developed a methodology for cybersecurity assessment using the National Framework for Cybersecurity and Data Protection [63]. This methodology introduces several innovative elements, structuring its activities into three phases: Contextualization, Measurement, Assessment.

- *Contextualization*: in this phase, the methodology selects and evaluates, in terms of priority and maturity, the subcategories of the National Framework of interest concerning the situation. This process is done through the combination of existing (or new ones) contextualization prototypes, all based on general informative references and those specific to one's sector.
- *Measurement*: at this stage, the distance between the current state and the target state is noted. One or more interviewers provide, using questionnaires formulated according to the target profile, to assess the level of achievement and implementation of the identified controls. The output of this phase is the current profile, which is a summary of the organization's security posture based on the controls and subcategories identified at the time of the assessment.
- *Assessment*: in the final phase, the results obtained in the previous phase can be read as an assessment of the distance between the current profile and the target profile. The result is embodied in a score of completion of the identified actions and an additional score that represents the degree of maturity with which the actions are implemented.

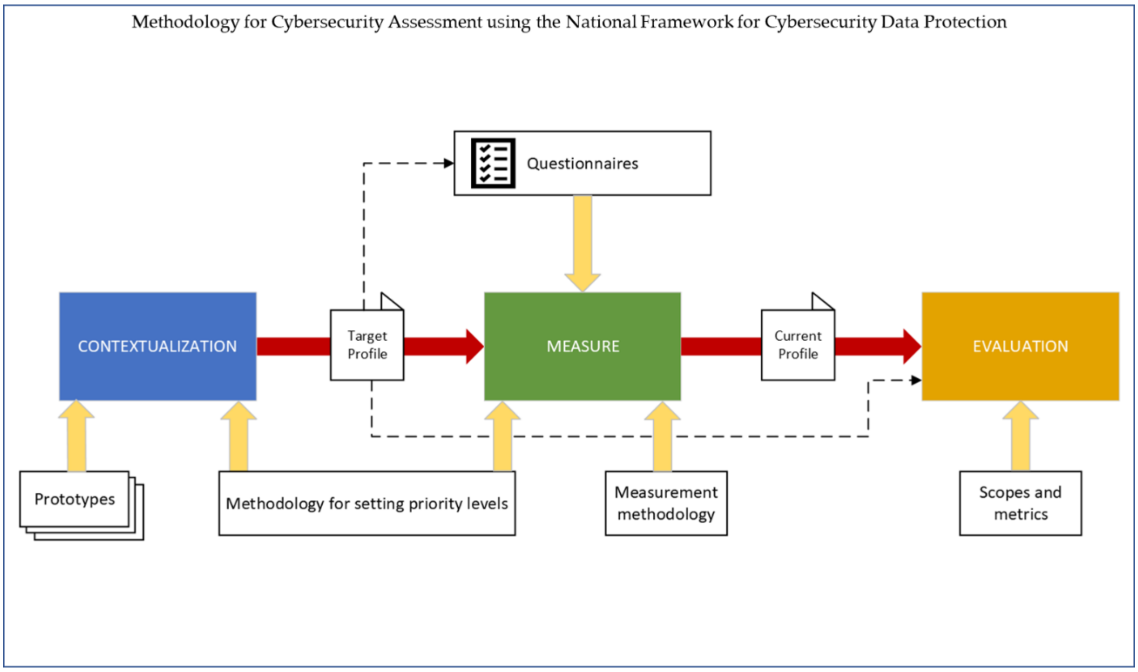

Methodology for Cybersecurity Assessment using the National Framework for Cybersecurity Data Protection
Questionnaires
CONTEXTUALIZATION
Target Profile
MEASURE
Current Profile
EVALUATION
Prototypes
Methodology for setting priority levels
Measurement methodology
Scopes and metrics

Figure 8 - Methodology for Cybersecurity Assessment using the National Framework for Cybersecurity Data Protection translated from [63]

2.5. Social Engineering

The term Social Engineering refers to a set of approaches in which attackers use human channels to achieve their goal. In this scenario, hacking attempts increasingly focus on the human vulnerabilities of an information elaboration system instead of lapses in software or hardware. [64] in 2011 pointed out that human has become the most vulnerable part of systems. His statement is confirmed today by the number of cyber incidents involving the human element. According to Verizon's 2022 Data Breaches Investigations Report [65], 82% of data breaches involved a human element. This includes incidents related to social attacks, errors, and misuse. IBM's Cost of a Data Breach Report 2022 [66] found that the two most expensive forms of data breach were the result of skill-based errors. The 2022 CLUSIT report [67] also puts a focus on cyber incidents related to human error. Indeed, it states how in just one year the percentage of small organizations (1-49 employees) that experienced incidents caused by employees grew from 25% to 32%. Particularly interesting is the fact that these companies mentioned "inattention" as one of the main causes: 46% of respondents emphasized the significant incidence of this factor in relation to the incidents that occurred. However, in most cases, "inattention" seems not to be the term that exhaustively describes the scenario, but this will be the subject of Chapter 4. The reported data also show how the pandemic environment has accelerated the growth of such cyber threats [68] in multiple domains.

One of the most challenging problems in social engineering analysis is its multidisciplinary, involving not just information security but also psychology and sociology. There is no agreement on its definition. To define social engineering, researchers from other domains are likely to adopt their own vocabulary.

From a technical point of view [69] their ontology for social engineering defines *"a social engineering attack as an attack that applies one or multiple social engineering attack techniques, targets one particular person who has at least one human vulnerability, and is performed by a social engineer through a particular type of attack media."* In particular, they refer to a *"human vulnerability as a reason that causes a high possibility for a human asset to take out actions as the intending results of social attacks initiated by social engineers."* Another notable definition refers to the work of [70] or which define social engineering as "the science of using social interaction as a means to persuade an individual or an organization to comply with a specific request from an attacker where either the social interaction, the persuasion or the request involves a computer-related entity". Despite researchers affirming there is not common consensus on the definition, they agree that the goal of the attack is to exploit the human layer of cyberspace to reach their final objective. There can be multiple reasons for conducting an attack: the most obvious is economic, political, and related to industrial espionage; but there are many documented cases of attacks carried out for fun, personal revenge, terrorism, activism, or collecting private information resulting in ransom demands or sale to third parties.

Nevertheless, it is interesting to understand how these types of attacks take place.

[71] represent in their work the "core" entities, concepts that significantly define or influence the domain of Social Engineering in Cybersecurity. The circular arrow depicted in Fig.1

indicates the typical attack process. The most common scenario involves defining the attacker and his or her motivations, acquiring the information needed to deploy an effective attack by exploiting vulnerabilities and evaluating the efficiency of the attack with respect to the intended objectives.

- Attacker

  The attacker is the entity that plans and/or carries out the attack. Characteristics of the attacker are:
    - The numerosity of the attacker: individual, group, or organization.
    - Position of the attacker in relation to the victim: internal or external.
    - Status of the attacker: a physical person or bot.

- Social Engineering Information

  The success of an attack depends primarily on the quantity and especially the quality of Social Engineering information available, namely: personal information of the targets (victims), information about the organization, information about the network, and information about social relationships. Any public information, leaked in cyberspace could provide attackers with important resources, to learn about the environment and context, discover targets, find vulnerable human factors, and cyber vulnerabilities, and useful details to formulate attack strategies specific to each victim.

- Attack strategy ad methods

  In terms of attack methods, the variables to be evaluated are available resources, environment, target, and related vulnerabilities. According to the literature, there are two different strategies for social engineering attacks: the "forward strategy" and the "reverse strategy" [71]. The first type, which is also the most widely used, involves a direct attack toward the target, aimed at penetrating its defenses; the "reverse," on the other hand, seeks to place the target in the condition of having to personally contact the attacker, since it is considered a legitimate, authoritative, and reliable source. At this stage, the objective of the attack is thus to establish a degree of trust between victim and attacker, and to obtain as much information as possible.

  The most exploited attack vectors in social engineering are:
    - SMS/voice call/e-mail
    - Sharing of infected hardware (e.g., USB pen drive).
    - In-person interaction.

  To:
    - Installing malware.
    - Obtaining confidential access keys.
    - Providing links to fraudulent websites.
    - Installing malicious applications.

  In terms of SE techniques to support the attack, it is worth mentioning:

  - Phishing / Spear Phishing
  - Pretexting
  - Baiting
  - Tailgating

- Target vulnerability and human factor

  The International Ergonomics Association [72] defines the study of the human factor as the scientific discipline concerned with understanding the interactions between human beings and the elements, thus highlighting the centrality of the user and, consequently, the factors involved in social interactions, such as devices (e.g., PCs, cell phones), tasks (e.g., internet browsing), responsibilities related to hierarchical role (e.g., knowledge of sensitive data) and environment (e.g., office, home).

  The target vulnerabilities of social engineering can be traced to four aspects [73]:

  - Cognitive and knowledge: e.g., ignorance, inexperience, bias, conformity, intuitive judgment, mental shortcuts.
  - Behavior and habit: e.g., laziness, negligence, fixed patterns of action, behavioral habits.
  - Emotion and feeling e.g., fear, curiosity, anger, excitement, tension, happiness, sadness, disgust, surprise, guilt, impulsiveness.
  - Human nature_ e.g., helpfulness, self-love, sympathy.
  - Personality traits (Extraversion, Agreeableness, Conscientiousness, Openness to Experience, Neuroticism [74]).
  - Individual characters: e.g., credulity, friendliness, courtesy, humility, envy.

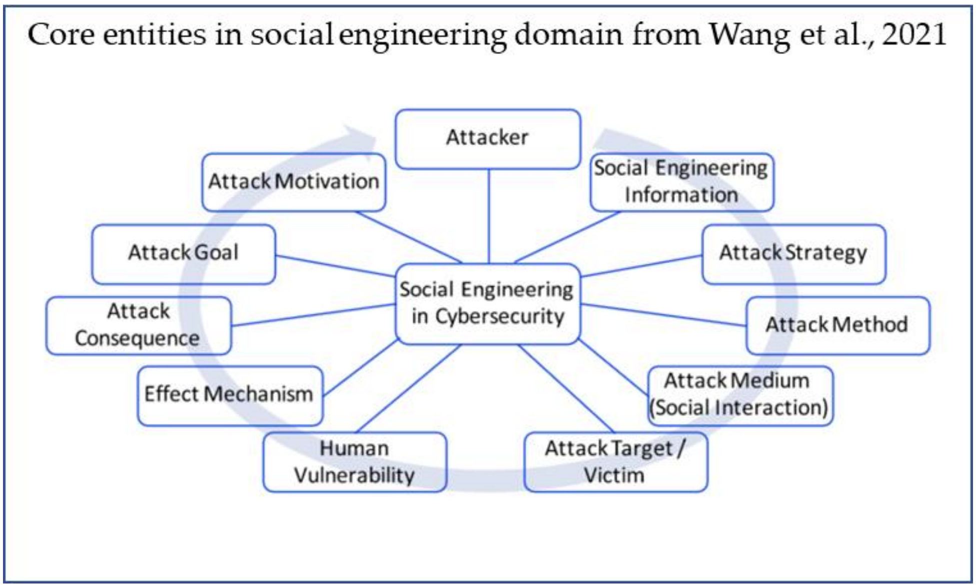

Core entities in social engineering domain from Wang et al., 2021
Attacker
Attack Motivation
Social Engineering Information
Attack Goal
Attack Strategy
Social Engineering in Cybersecurity
Attack Consequence
Attack Method
Effect Mechanism
Attack Medium (Social Interaction)
Human Vulnerability
Attack Target / Victim

Figure 9 - Core entities in social engineering domain retrieved from [71]

Instead, the work of [75] emphasizes how social engineering (SE) is not limited to the interaction between attacker and victim but, considering the complex process of action, involves the centrality of the role of the organization and security policies. This multidimensional approach has changed the original view of this type of attack by placing at its core the need for systems thinking that encompasses the entire organization. [76]'s work echoes this consideration, analyzing how personality traits influence vulnerability to SE attacks by bringing several situations that occur in the business context as examples. The interesting result that emerges is how the characteristics of victims' personality traits can support and guide the corporate cybersecurity process. In fact, an interesting result that this thesis will assume later is how individual factors of personality traits relate to the success or failure of SE attacks. Highlighting the need to adapt for instance mitigation strategies according to each personality and level of insiderness or propose customized training programs based on profiles.

3. Cyber resilience opportunities and challenges

Given the wide range of application domains, the flexibility of cyber-physical systems, and the different threats involve, this PhD journey started with a general exploration of cyber-related domains of application. The method used applied natural language processing (NLP) techniques and a k-means clustering algorithm on article metadata to identify clusters of topics in the cyber resilience research field. Moreover, NLP helped in understanding the domains of application of cyber resilience, the possible threats, and the related effects mentioned in the literature. The application of these techniques helped in the definition of the first research question of the thesis.

After this exploration, the thesis continued with a more focused investigation of the resilience dimensions involved in the study of CPSs using the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.

The following paragraphs introduces the techniques used and presents the results of the analysis using NLP. Then, the systematic literature review published in Computers and Industrial Engineering in 2021 is appended.

3.1. Cyber Resilience : a cross-domains topic exploration

3.1.1. Natural Language Processing Techniques and K-Means Clustering Modeling

NLP is an artificial intelligence component that is used to help computers understand human natural language. These techniques are part of a field of study in which computer science, artificial intelligence, and linguistics are combined, resulting in a programming method aimed at processing large amounts of data with various objectives, ranging from information extraction via various text processing methods to emotion detection, i.e., Sentimental Analysis [77], [78]. Using algorithms and analytical approaches, NLP enables retrieval, analysis, and information-condensing processes, as well as detecting models and patterns, labeling massive datasets, and expediting procedures for the presentation of hidden structures. In this setting, printed text data is an important input. Text mining techniques are becoming more popular, and there is also a rising interest in the industrial sector. NLP has been used to manage industrial risks, for example. While NLP was utilized in the study provided by [79] to analyze incident and accident narratives, text mining was employed in the work of [80] to extract additional information from narratives in reports linked to natural gas distribution pipes. NLP has also been used to enhance complicated production, such as semiconductor manufacturing, by evaluating documents related to the manufacturing process and incorporating lessons gained ([81]. [82]) have presented a method for investigating technical trends in smart manufacturing by analyzing language included in patent data using NLP and topic modeling algorithms. Similarly, [83] employed text-mining algorithms to uncover key digital technologies explored in the mining sector, as well as relationships between major digital trends.

For the sake of this thesis project, text analysis logic was utilized to extract concepts, areas of exploration, and research topics from a vast number of documents acquired from the existing literature [84], [85].

The objective was to read, understand, and make sense of the text in a valuable manner. Syntactic analysis and semantic analysis were the main techniques used to process the text. The proposed approach can be divided into four main steps: paper collection, pre-processing phase, text-feature extraction, and clustering modeling.

1. Paper collection

   The initial step is collecting a set of research articles gathered from large databases of peer-reviewed literature. The approach is applied to articles' titles, keywords, and abstracts. Before going to the following step, any duplicate article must be removed.

2. Pre-processing

   In this second step, a set of data pre-processing activities is conducted. Data pre-processing is a method of data mining that involves transforming raw data into a reasonable format for topic modeling and cluster analysis. The pre-processing stage is developed using NLTK library in Python. For the analysis of text data collected the following natural language processing steps must be pursued.

   - Tokenization

     Since documents are unstructured information, they must be divided into linguistic units. The process of splitting a phrase, sentence, paragraph, or entire text document into smaller units is called tokenization and these pieces are called tokens. There are different tokenization strategies, in this article white spaces between words are considered a separator. In this step also punctuations are removed.

   - Sparse terms and stop-words removal

     Stop-words are commonly used words (e.g., articles, prepositions, pronouns, etc.) that occur frequently in article corpora but do not carry any meaning on their own. Depending on the domain and language there will be a distinct set of stop-word. Moreover, for better performance, it is recommended to remove search terms (e.g., terms used in the query to collect the articles) that may bias the clustering results. Sparse terms are instead terms that occur in less than 1% of the documents. All these terms are removed from the article corpora in this step.

   - Stemming / Lemmatization

     Stemming is a process of reducing words to their root form. Lemmatization instead is a more complex version of stemming that uses part-of-speech (POS) tagging for each word, (e.g., verb, noun, adverb, etc). First, each token is tagged with a POS, second lemmatization applies a different stemming rule to each token depending on the tagged POS. This allows diminishing the lexical sparsity of the corpus since two words that have the same root result in the same output.

- Ngrams

  N-grams are n words frequently occurring together in the document. N-grams estimate the probability of the next item in a word sequence calculating the occurrence of the next word with respect to the previous one. In this analysis are consider b-grams two-word sequence of words – and tri-grams – a three-word sequence of words.

3. Text Feature Extraction

The last step in the natural language processing pipeline is TFIDF vectorization. The TFIDF measure reflects how important a word is to a document in a collection of documents and it consists of two parts. The first part is term frequency which counts how many times a word appeared in each document considered. The second part is inverse document frequency which is responsible for reducing the weights of words that occur frequently and increasing the weights of words that occur rarely [86].

$$TFIDF(Word, Document, AllDocuments) \\ = TermFrequency(Word, Document) \\ * InverseDocumentFrequency(Word, AllDocuments)$$

4. Clustering

The approach proposed the K-Means algorithm. K-means is the most used unsupervised clustering algorithm and partitions the N documents in K disjoint clusters, defining as K the optimum number of clusters. The objective of the K-Means algorithm is to divide a given number of samples into a deliberately selected number of clusters. Once the number of clusters is set, the algorithm randomly selects k samples for the centroids. Then each observation is assigned to a cluster to minimize the within-cluster sum of squares. Next, the mean of the clustered observations is calculated and used as the new cluster centroid. Then, observations are reassigned to clusters, and centroids are recalculated in an iterative process until the algorithm reaches convergence. The optimum number of clusters is chosen using the silhouette score [85]. The silhouette is a measure of how close each point in one cluster is to points in the neighboring clusters. Given a cluster *Ci*:

$$s(i) = \frac{b(i) - a(i)}{max\ \{a(i), b(i)\}}\ se\ |C_i| > 1 s(i) = 0\ se\ |C_i| = 1$$

where *b(i)* is the minimum average distance of I to all points in any other cluster, of which *i* is not a member. *a(i)* is the average distance between *i* and all other data points in the same cluster. *a(i)* is therefore a measure of how well *i* is assigned to its cluster (the lower the value, the better the assignment). The coefficient *s(i)* varies between -1 and 1. Whereas, a value close to -1 means that the value is assigned to the wrong cluster [87].

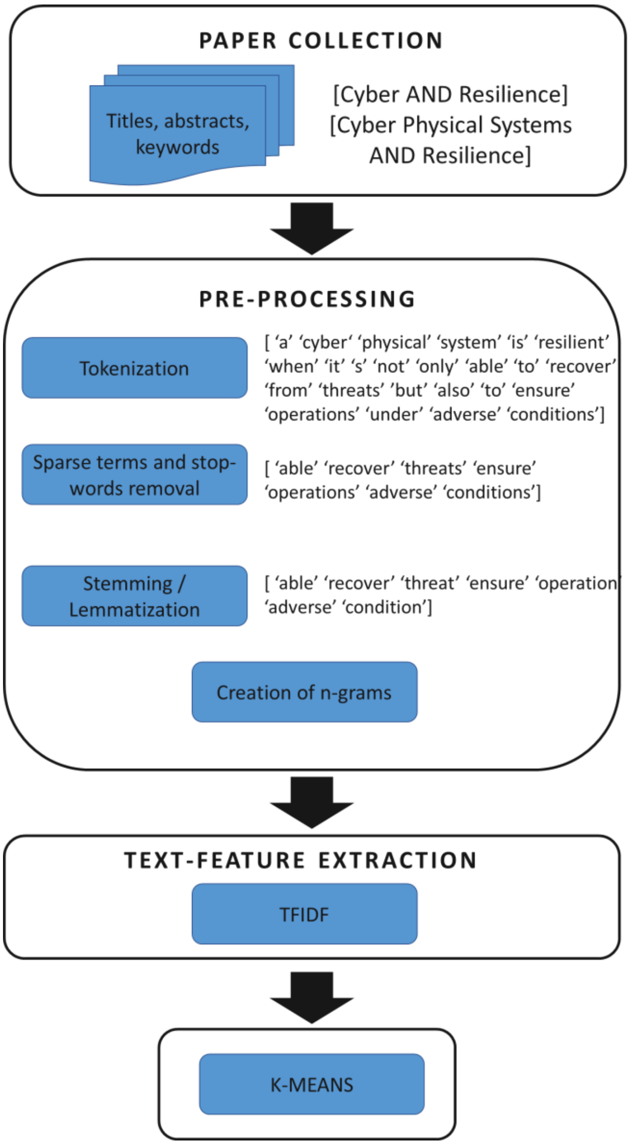

PAPER COLLECTION
Titles, abstracts, keywords
[Cyber AND Resilience]
[Cyber Physical Systems AND Resilience]
PRE-PROCESSING
Tokenization
['a' 'cyber' 'physical' 'system' 'is' 'resilient' 'when' 'it' 's' 'not' 'only' 'able' 'to' 'recover' 'from' 'threats' 'but' 'also' 'to' 'ensure' 'operations' 'under' 'adverse' 'conditions']
Sparse terms and stop-words removal
['able' 'recover' 'threats' 'ensure' 'operations' 'adverse' 'conditions']
Stemming / Lemmatization
['able' 'recover' 'threat' 'ensure' 'operation' 'adverse' 'condition']
Creation of n-grams
TEXT-FEATURE EXTRACTION
TFIDF
K-MEANS

Figure 10–- Steps for text processing

3.1.2. Results

The process just described was repeated twice. The first time, an overly broad search key was chosen: Cyber AND Resilience, which collected 1623 results. These results were preprocessed and clustered reporting the clusters shown in Figure 11.

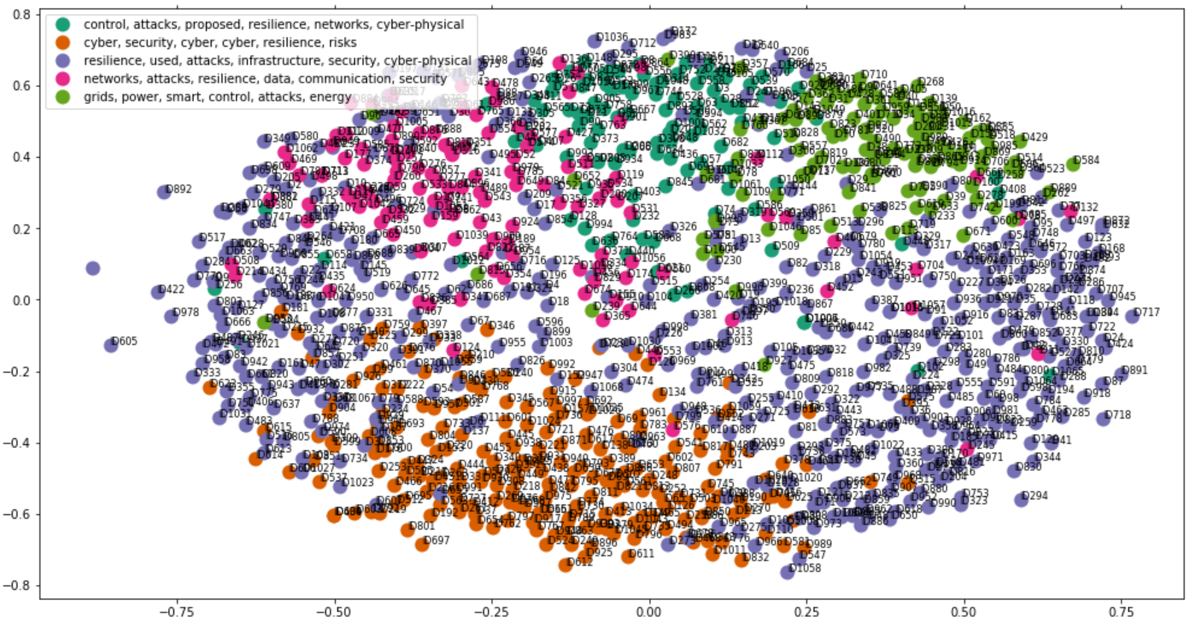
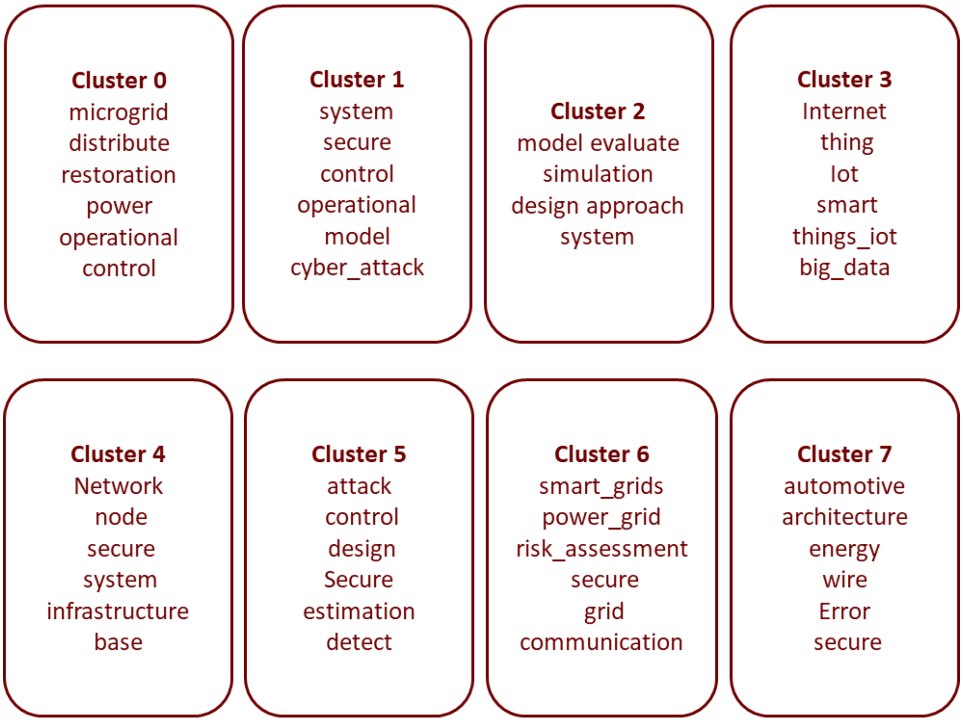


Figure 11–- Cluster analysis using the search key "Cyber AND Resilience"

The highlighted clusters demonstrate how the topic is not only interdisciplinary but also can be poorly classified into distinct areas of research. This phenomenon is related to the variety of topics that are encompassed in the major theme of cyber resilience. Indeed, this encompasses topics such as systems security, the ability to detect risks, control networks and critical infrastructures, across different domains: from communications to the energy sector. It was thus chosen to repeat the analysis by going to narrow the search key to Cyber-Physical Systems AND Resilience.

A CPS is a computer system able to interact continuously with the physical system in which it operates. They are complex and interconnected systems integrated into our everyday lives forming the basis of smart infrastructures, products, and services. They have been applied in various fields, including energy, healthcare, manufacturing, transportation, and smart environments. These systems enable the generation and acquisition of data and support decision-making by ensuring reliable and secure operations in infrastructures. However, the physical world in which CPSs operates is not entirely predictable and typically affected by multiple risks such as environmental risks or cyber-attacks. Considering the industrial nature of this Ph.D. and the increasing presence of CPSs in factories and their application in all sectors, the focus of the review was limited to CPSs. Figure 12 reports the cluster that resulted from the analysis.



Cluster 0
microgrid
distribute
restoration
power
operational
control

Cluster 1
system
secure
control
operational
model
cyber_attack

Cluster 2
model evaluate
simulation
design approach
system

Cluster 3
Internet
thing
Iot
smart
things_iot
big_data

Cluster 4
Network
node
secure
system
infrastructure
base

Cluster 5
attack
control
design
Secure
estimation
detect

Cluster 6
smart_grids
power_grid
risk_assessment
secure
grid
communication

Cluster 7
automotive
architecture
energy
wire
Error
secure

Figure 12–- Cluster analysis using the search key "Cyber Physical System AND Resilience"

What emerged clearly was the diffusion of the topic across multiple domains underling the relevance of the topic. Moreover, this first analysis showed that while many documents deal with this kind of problem for systems including CPSs, there was still a lack of a cross-domain review on the different dimensions of system's resilience, explored considering CPSs from a joint technical and socio-technical research dimension. A more detailed exploration of the full papers belonging to the clusters enabled the definition of the RQs underlying the more extensive literature review that was then conducted.

Below the research questions defined are summarized:

1. To which extent do CPSs contribute to the resilience of technical and socio-technical systems?
2. Which are the approaches available in the literature to understand, measure, and model the resilience for those systems?

3.2. Cyber Physical Systems Resilience: A Systematic Literature Review

As previously mentioned, once the broad area of cyber resilience was explored, a more focused extensive literature review has started. The scope was to survey available literature for understanding to which extent CPSs contribute to system resilience, and to synthetize the approaches developed in this domain. More than 500 documents were reviewed through a protocol based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) review technique. This survey identified main models and methods categorizing them based on the hazards of interest and their effects on security, privacy, safety, and business continuity. It also summarizes main conceptual frameworks and metrics used to assess and compare the resilience capabilities of a system including also CPSs.

The cross-domain review answered the first research question of this thesis by emphasizing that the available approaches to increasing resilience are primarily techno-centric while still highlighting emerging trends toward more systemic representations of threats even to socio-technical systems.

Moreover, the survey identified important research gaps and objectives which have been the baseline for the research direction followed in the second part of the thesis. Research gaps that emerged can be summarized as follows:

1. **The need for human-in-the-loop in CPS: humans as part of the solution.**
   First, the survey stressed the importance of humans. The analysis of the resilience of large socio-technical systems of course includes humans. Traditionally, the reviewed contributions identify human users as one of the major threats in today's technologies. Only a few studies started considering humans to improve the effectiveness of CPSs resilience solutions. The idea beyond this is the need to structure CPSs in a way that ensures people are incorporated as part of its process, defining a more human-centric cybersecurity.
2. **The shift toward cyber-socio-technical systems: integrating cyber security & cyber safety**
   In line with the first research gap, the review showed a limited number of resilience metrics that consider both technical and socio-technical aspects. Therefore, it emerges the need for a framework of flexible socio-technical metrics to evaluate the resilience of systems including CPSs. Related to models and methods, promising is the use of system theoretic models such as STAMP and STPA-SEC. The applications of STAMP and STPA-SEC is interesting in the safety of CPS. These models can help at integrating safety and security needs. The reason is that STAMP differs from other cybersecurity approaches in the way it maps interdependencies between technical and human elements representing a promising research area for future socio-technical cyber analyses. There is still little knowledge about the safety-related consequences of systems under cyber-attacks and the importance of integrating cyber security and cyber safety.
3. **Building resilience in cybersecurity.**
   Modern systems face new threats and there is a dominance of cyber-attacks to CPS. However, these hazards are constantly changing, forcing a continuous reconsideration of strategies to ensure operational continuity. Future studies should keep ensuring that the cyber defense mechanism is dynamic and real-time. Threats and incidents become more sophisticated, and it is no longer possible to fight them in small-world scenarios, managing variability rather than simply trying to reduce it.
4. **System modeling and simulation.**
   Future studies should encompass a greater use of DT and System dynamics. Compared to their widespread usage in other industrial settings in this context, the usage is limited which on the contrary may constitute a valuable research method to

monitor CPSs and evaluate the evolution of system performance in the case of an attack.

For details of the research carried out, and for a complete account of the articles collected from both a bibliometric and a described perspective, the article *"Discussing resilience in the context of cyber physical systems"* published in *"Computers & Industrial Engineering"* by authors Silvia Colabianchi, Francesco Costantino, Giulio Di Gravio, Fabio Nonino, and Riccardo Patriarca is attached below.

3.2.1. Appended Paper 1: Discussing Resilience in the Context of Cyber Physical Systems

Computers & Industrial Engineering 160 (2021) 107534

Contents lists available at ScienceDirect

Computers & Industrial Engineering

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

Review

Discussing resilience in the context of cyber physical systems

Silvia Colabianchi [a,*], Francesco Costantino [a], Giulio Di Gravio [a], Fabio Nonino [b], Riccardo Patriarca [a]

[a] *Dept. of Mechanical and Aerospace Engineering, Sapienza University of Rome, Via Eudossiana 18, 00184 Rome, Italy*
[b] *Dept. of Computer, Control and Management Engineering, Sapienza University of Rome, Via Ariosto 25, 00185 Rome, Italy*

ARTICLE INFO

Keywords:
Safety
Security
Business continuity
Cybersecurity
Socio-technical systems
Cyber resilience

ABSTRACT

Cyber-Physical Systems (CPSs) are increasingly more complex and integrated into our everyday lives forming the basis of smart infrastructures, products, and services. Consequently, there is a greater need for their ability to perform their required functions under expected and unexpected adverse events. Moreover, the multitude of threats and their rapid evolution pushes the development of approaches that go beyond pure technical reliability, rather encompassing multi-dimensional performance of a socio-technical system. These dimensions call for the notion of resilience, to be used as a staging area for modelling system performance. While a large number of documents deal with this kind of problem for systems including CPSs, a comprehensive review on the topic is still lacking. The scope of this paper is to survey available literature for understanding to which extent CPSs contribute to system resilience, and to synthetize the approaches developed in this domain. More than 500 documents were reviewed through a protocol based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) review technique. This survey identifies main models and methods categorizing them on the basis of the hazards of interest and their effects on security, privacy, safety and business continuity. It also summarizes main conceptual frameworks and metrics used to assess and compare the resilience capabilities of a system including also CPSs. This cross-domain survey highlights the dominant techno-centric unit of analysis for available literature, still highlighting emerging trends towards more systemic representations of system threats, even socio-technically oriented, and respective modern investigation approaches.

1. Introduction

A Cyber-Physical System (CPS) is a computer system able to interact continuously with the physical system in which it operates. CPSs have been applied in various fields from smart grids to medical devices, smart factories, intelligent transport systems, smart cities, and smart buildings, all these systems are today known as Cyber-Physical Systems (Gujrati, Zhu, & Singh, 2015; Xu, 2020), These systems enable the generation and acquisition of data and support decision making by ensuring reliable and secure operations (Mörth, 2020). However, the world in which CPSs operates is not entirely predictable and typically affected by multiple risks such as natural hazards or cyber-attacks (Khalid, Rehman, & Shafique, 2020). On this regard, Mokalled (2019) affirms the need to consider not only the cyber risks among the threats but also additional non-IT threats which could affect the physical assets that are necessary for the operation of the CPS. They classify the threats for CPSs into human errors, third-party failures, malicious actions, system failures, and natural phenomena. Björck (2015) instead divides the adverse cyber events into two main groups. The first group is defined as *"Acts of God"* which comprise events caused by nature, and *"Acts of man"* in which events caused by people, intentional or unintentional, are considered.

Given the multitude of threats and their rapid evolution and the ever-increasing dependence on CPSs of modern systems, there is now a demand to provide approaches that go beyond pure technical reliability, encompassing the notion of resilience. A system is resilient when it is not only able to recover from threats, but also to ensure operations under adverse conditions (Patriarca, 2018). It is indeed no longer possible to prevent and predict all the possible threats that these socio-technical systems suffer. The concept of resilience might help in treating adverse events as part of CPSs normal operation (Mailloux & Grimaila, 2018). However, when we focus on these new vulnerabilities, and consequently on the topic of cyber-resilience, this awareness is not yet fully developed (Björck, 2015; Jacobs, Hossain-Mckenzie, & Vugrin, 2018). The topic of system resilience, extended considering CPSs, is

[*] Corresponding author.
E-mail address: silvia.colabianchi@uniroma1.it (S. Colabianchi).

https://doi.org/10.1016/j.cie.2021.107534
Received 28 November 2020; Received in revised form 19 April 2021; Accepted 4 July 2021
Available online 6 July 2021

widely studied and investigated through multiple approaches. CPSs are defined by the National Institute of Standards and Technology (NIST) as *"hybrid networked cyber and engineered physical elements co-designed to create adaptive and predictive systems for enhanced performance. Performance metrics include safety and security, reliability, agility and stability, efficiency and sustainability, privacy"* (Institute of Standards, 2014). Bennaceur (2019) specify the need of designing resilient CPS in order to achieve the performance metrics specified by the NIST. The authors affirm that a CPS is resilient if it is able to self-adapt to deal with change. They also stress the need to rethink methodologies, tools and models to provide guarantees about the overall behaviour of CPS.

Several reviews have been proposed: some were focused on the domain of application, others focused on organizational aspects or specific approaches to resilience analysis. Starting from studies focused on CPS (Boyes, 2013; Bennaceur, 2019; Dibaji, 2019; Weerakkody, 2019) or IoT (Cheng, 2018; Schoitsch, 2018; Sun & Yang, 2018; Ratasich, 2019) it emerges a need for research to achieve trustworthy sensors. Boyes (2013), Lin (2017) and Schoitsch (2018) provide an overview of technological and societal aspects as well as emerging threats linked to these systems. Weerakkody (2019) investigate the challenges of achieving reliable control and resilient operation in CPS and Dibaji (2019) carry out a survey of systems and control methods proposed for the security of CPS. Security and privacy issues of these systems are also addressed in Ratasich (2019) and Gupta et al. (2020). In Ratasich (2019), the authors summarize the state of the art of existing work on anomaly detection, fault-tolerance, and self-healing methods applicable to achieve IoT resilience. In Gupta et al. (2020) are instead presented tools and challenges related to privacy protection in CPS using AI. Bennaceur (2019) presents new challenges in CPS and identifies research challenges when modelling and engineering CPS, stressing the importance of designing resilient CPS. Several surveys have been proposed for the energy domain. Li, Shahidehpour, and Aminifar (2017) present the application of cybersecurity to the control of distributed power systems; Arghandeh (2016), Haggi (2019), Inderwildi (2020) and Mohebbi (2020) focus on the application of smart grid technologies and CPS to enhance the resilience of the power systems; (Das, 2020) propose a review and analysis of qualitative frameworks and quantitative metrics for studying the resilience of the smart grid; Arghandeh (2016) review the state of the art of reliability modelling methods and the evaluation indexes of urban multi-energy systems. The relevance of CPS and their resilience emerges also from numerous research studies focused on production systems in the emerging scenario of industry 4.0. For example, a survey on the security control and attack detection for industrial CPS is proposed by Ding (2018). A set of diverse definitions of resilience and robustness and reflection related to new emerging technologies involved in the industry 4.0 and future challenges are proposed in (Lee, Bagheri, and Kao (2015), Lee, Bagheri, and Jin (2016), Maurer and Schumacher (2018), Panetto (2019), and Tao (2019). Finally, Wu, Goepp, and Siadat (2019) and Lins and Oliveira (2020) classify the current research activities within Cyber Physical Production Systems (CPPSs) with a special focus on design and implementation approaches. Other reviews involved the resilience of the automotive industry (Fraga-Lamas & Fernández-Caramés, 2019), the security of the digitally managed water distribution systems (Gupta et al., 2020; Mohebbi et al., 2020), a survey of cyber risk in supply chains (Ghadge, 2019), and cyber-attacks towards airports (Lykou, Moustakas, & Gritzalis, 2020).

Nonetheless, as observable from the previous paragraphs, there is still a lack of a cross-domain review on the different dimensions of system's resilience, explored in light of CPSs from a joint technical and socio-technical research dimension. Due to the relevance of CPS in a profusion of modern industrial settings, the scope of this review is expected to gather approaches available in multiple settings to favour a cross-domain, more general, learning approach. More formally, the research questions of this paper can be summarized as follow:

– To which extent do CPSs contribute to the resilience of technical and socio-technical systems?
– Which are the approaches available in the literature to understand, measure, and model the resilience for those systems?

To answer these questions a scoping review is developed, adopting a systematic approach for defining eligible studies, in line with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines (Moher, 2009).

The remainder of the paper is organized as follows. Section 2 details the database and query used for the analysis and the systematic approach followed for the review. Section 3 presents a bibliometric analysis of the dataset of documents considered for review. Section 4 details a synthesis of main interpretative findings, where articles are described according to the approaches used to understand, measure, and model system's resilience, also including CPSs. Finally, Section 5 discusses the outcomes of the research, offers critical reflections on the field and possible future research paths. The conclusion summarizes the review and underlines the open research questions.

2. Materials and methods

This review investigates articles and conference papers following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, given in (Moher, 2009). PRISMA guidelines define a systematic process of study identification, screening, eligibility, and inclusion. The detailed process followed for this study is described in the workflow shown in Fig. 1.

2.1. Identification

The review is conducted through a search in Scopus database, of contributions indexed up to 31st March 2020. The choice of this database took into consideration its relevance in the academia. Scopus is a leading source, with over 75 m records across 25,000 journals, sourced from more than 5000 publishers (RELX, 2019). It indexes several well-regarded journals and conferences besides being used in other bibliographic reviews concerning complex systems (Patriarca, 2020), or even cyber-resilience (Mouelhi, 2019).

The first step of the review defines the scope of the search query. The search query looks for every paper that uses "cyber-physical system" to refer to the systems considered in association with "resilience" to broadly collect all the articles that explore at least one dimension of resilience in relation to CPSs. However, for the purpose of having broader inclusion criteria, quotation marks are used to ensure that terms composed of multiple words are searched together, as well as asterisks are used to include both singular and plurals terms, and derived terms.

The scope of the research is limited to journal and conference articles, published in English excluding articles classified as reviews, as well as book chapters, books, notes, erratum.

In summary, the search query for the database is:

(TITLE-ABS (resilien* AND cyber*)) AND (TITLE-ABS-KEY ("cyber physical system" AND resilien*)) AND (LIMIT-TO (SRCTYPE, "p") OR LIMIT-TO (SRCTYPE, "j")) AND (LIMIT-TO (DOCTYPE, "cp") OR LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English"))

504 articles are identified through this search query. Then, 10 extended abstracts, 4 panels, and 13 articles for which the full text is not available are excluded. Therefore, the identification phase is concluded with 477 articles selected.

2.2. Screening

In the screening phase, each article is screened in the title, abstract, and keywords to evaluate if its research is adherent to the objective of the review. Among 477 articles, 33 are excluded since they do not consider cyber-physical systems or resilience, 5 are instead discarded

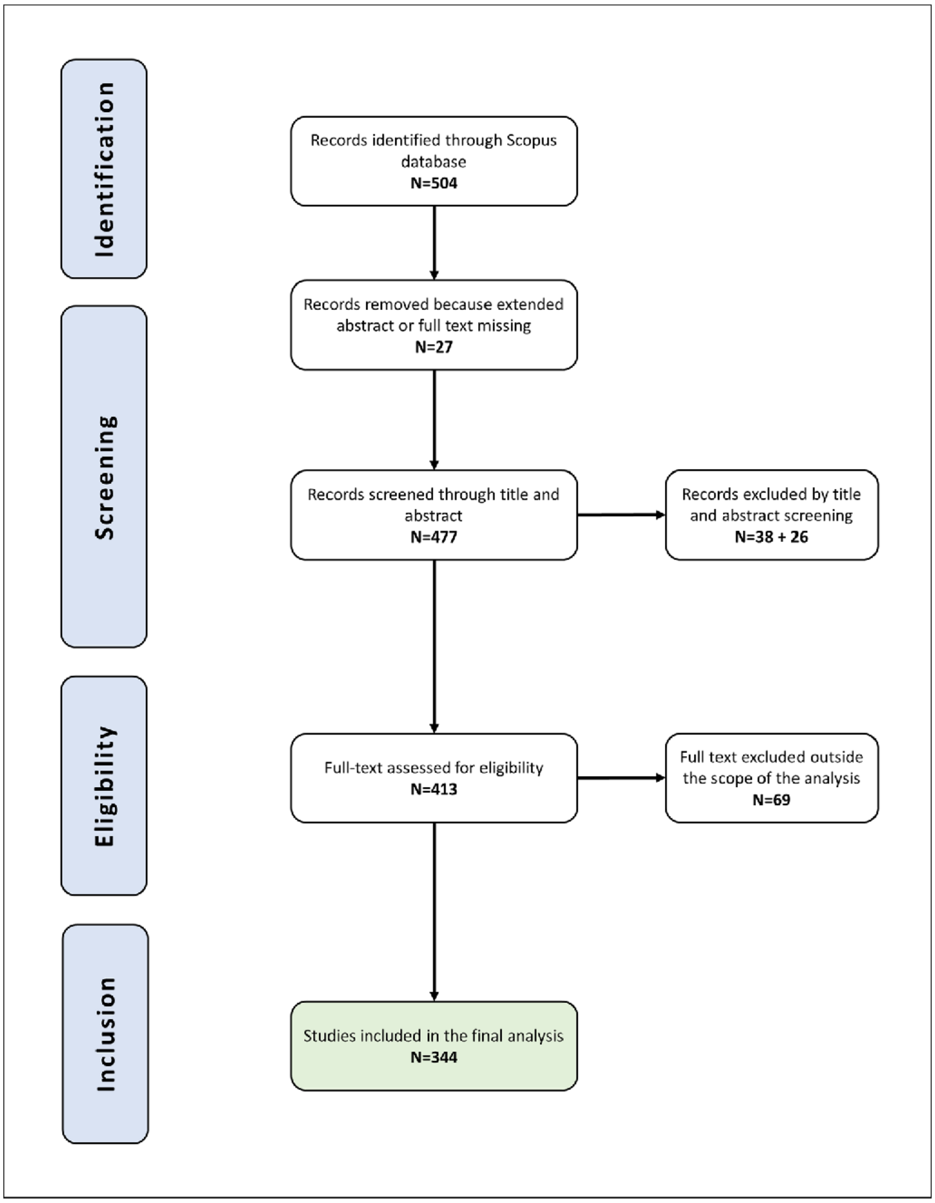

Identification
Screening
Eligibility
Inclusion
Records identified through Scopus database N=504
Records removed because extended abstract or full text missing N=27
Records screened through title and abstract N=477
Records excluded by title and abstract screening N=38 + 26
Full-text assessed for eligibility N=413
Full text excluded outside the scope of the analysis N=69
Studies included in the final analysis N=344

Fig. 1. Literature search strategy.

because their research concerns “error resilience”, which refers to the specific ability of a system to recover from errors in the results of computations (Abraham, Banerjee, & Chatterjee, 2017). Finally, 26 reviews are identified and excluded. These documents are not indexed as reviews in Scopus, and therefore they were not excluded in the initial search query.

2.3. *Eligibility*

During this phase, 413 articles are analysed. The research team reviewed the full texts and 69 were rejected as they do not meet the inclusion criteria.

2.4. *Inclusion*

A total of 344 articles are included for further analysis. These documents are reviewed thoroughly, to allow data extraction and to synthesize the information pertinent to the scope of this review. Data are organized using a framework that helps to categorize information concerning citation information, abstract and keywords, domains of application, hazard, hazard effects, resilience dimensions, and information concerning models, methods, and frameworks used in the papers. The fields reported in the framework are the following:

- Authors
- Title of the paper
- Year
- Source Title
- Number of citations
- DOI
- Abstract
- Keywords
- Document Type
- Domain

- Hazard
  - o Generic cyber attack
  - o Generic fault
  - o Generic disruption
  - o Generic natural hazard
- Hazard effect
  - o Safety
  - o Security
  - o Privacy
  - o Business continuity
- Resilience dimension
  - o Monitoring / Detection
  - o Mitigating
  - o Restoration
- Model
- Framework
- Method
- Metrics

3. Bibliometric analysis

The bibliometric analysis described in this section identifies the quantitative aspects of the research sector. Five primary aspects are analysed: source and document type, domain of application and their evolution over time, hazards and respective effects, the resilience dimensions, and the type of the approach (model, method, framework, metrics).

3.1. *Source and document type*

As shown by the pie chart in Fig. 2 there is a fairly balanced ratio between journal articles and conference proceedings, even though conference papers constitute approximately 58% of the total. Concerning the evolution over time of the publications, it is possible to observe in Fig. 3 a growing trend in the number of conference papers and articles related to CPS and resilience. This result is certainly linked to the spread of the technology for CPS, but also to an increasing interest in understanding the capacity to monitor, mitigate and respond to modern threats.

Subsequently, the analysis of the dataset shows the most relevant journals and conferences. 143 articles appear in 84 journals mainly related to the IEEE database.

On the other hand, 201 conference papers have been published in 160 conference proceedings, proving quite disperse results when coming to international events.

3.2. *Application domain*

Cyber-physical systems are today studied and used in multiple domains of application (Bou-Harb, Kaisar, & Austin, 2017). However, this review shows a fairly limited group of dominant domains in which resilience has been discussed. As shown in Fig. 4. Energy is by far the domain in which the resilience of CPS is investigated the most. Another domain investigated refers to “Digital Service Infrastructure” which includes studies focused on server, software, hardware, network, cloud, fog, and edge computing. “Energy” and “Digital Service Infrastructure” are also the domains that have started focusing on the research topic before all the others with the first documents published in 2011. Manufacturing systems are also a topic of wide interest reporting 32 documents: among them, 3 focuses on industrial IoT, 7 on CPPS, and 22 on manufacturing plants in general. Transportation domains are also investigated with documents mainly focused on Automotive and Aviation. Other relevant domains are “Smart Cities”, “Water”, which includes water treatment and distribution systems, “Defence”, “Satellite”, “Smart Home - Smart Workplace”. In Fig. 4 domains accounting for less than 5 documents each have been included under the label “Other”, i.e. “Healthcare”, “Nuclear”, “Supply Chain”, “Emergency Management”, “Oil & Gas”, “Nuclear”. Moreover, it is worth noticing a proposed domain labelled “Not specialized”. This category includes 104 documents that offer an approach to CPS and sensors in general, without any discussion in a specific domain. Finally, observing Fig. 5 Energy and CPS remain topics of great interest even showing a peak of publications from 2016. Moreover, it is also interesting to observe how there is a growing trend of documents in all areas starting from 2017.

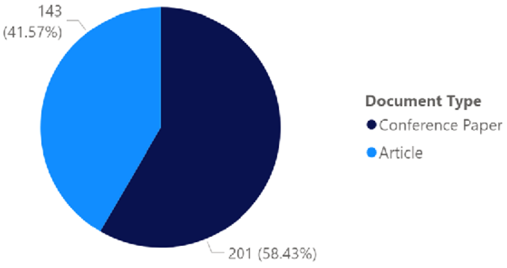


Fig. 2. Document type.

3.3. *Hazard*

In this section, the main hazards presented in the revised documents are analysed. Fig. 6 represents four major hazard categories: “Generic Cyber Attack”, “Generic Fault”, “Generic Natural Hazard” and “Generic Disruption”. As expected, cyber-attacks are by far the most studied threats in the field of cyber-physical systems. For this purpose, Fig. 7 shows a more detailed classification of this emerging threat for modern CPSs. Among all the cyber-attacks, Denial of Services Attacks, False Data Injection, Stealthy and Malware attacks are the most studied. Those documents that do not specify the cyber-attack investigated are labelled as “Various Cyber Attack” in Fig. 7. Moreover, Table 1 summarizes the most diffuse modes of attack against CPSs, the ways attacks occur and recent studies in which the resilience of CPSs to these attacks have been investigated.

3.4. *Hazard effects*

The analysis of the documents collected also identifies which effects the studies address. Fig. 8 shows the predominant role of security (approximately 60%) and business continuity (approximately 32%). Safety instead reports only 8% of the contributions. However, as can be seen in Fig. 8 there is a growing interest in safety in the most recent years (from a few documents before 2016 to 9 documents in 2018, 10 in 2019, and 3 in just the first three months of 2020).

3.5. *Resilience dimension*

As previously stated, all studies included in this review propose an approach to assess the resilience of cyber-physical systems. Before proceeding in the next paragraph with the analysis and discussion of these approaches it is possible to observe how there is greater attention to resilience measures related to the dimension of monitoring and mitigating (see Fig. 9). Specifically, 129 documents propose solutions to monitor and detect a fault or threat; 234, the majority among the studies analysed, address mitigation, fault tolerance, and vulnerability. Finally, 45 documents focus on recovering.

4. Descriptive findings

This section outlines some descriptive findings on the articles included in the final review dataset. In line with the framework structure, this section presents empirical contributions according to the type of approach used in the document. To facilitate a cross-domain

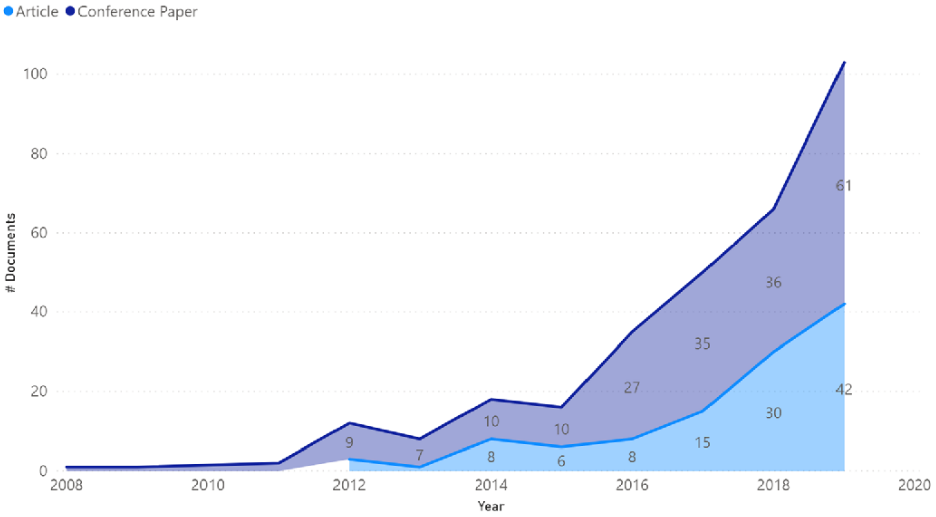
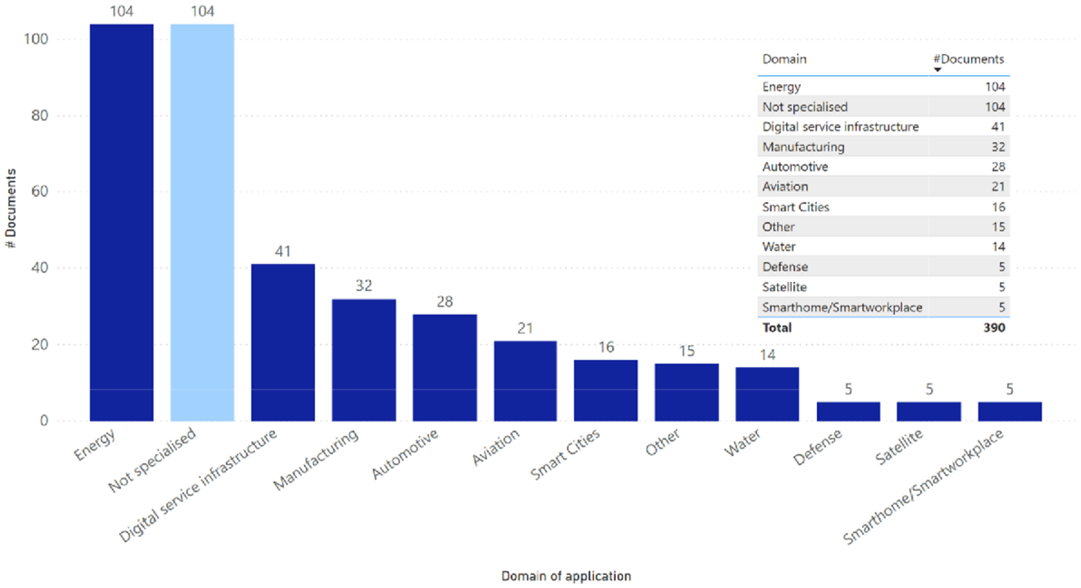


Fig. 3. Evolution of documents types over years.



Domain	#Documents
Energy	104
Not specialised	104
Digital service infrastructure	41
Manufacturing	32
Automotive	28
Aviation	21
Smart Cities	16
Other	15
Water	14
Defense	5
Satellite	5
Smarthome/Smartworkplace	5
Total	**390**

Fig. 4. Number of documents by domain of application.

understanding, further distinctions are made on the specific approach used, the key features of the domain of application, and the type of strategy to be implemented on the system. In those cases, where a document adopts one or more perspectives, or deals with several domains, it has been discussed under the category that represents the study.

Descriptive findings are described starting from the analysis of the simulation models, game theory models, graph theory, conceptual frameworks and resilience metrics. Subsequently, specific approaches and algorithms most used to analyse the resilience of CPS are presented.

4.1. Simulation

The resilience of CPS is evaluated here through simulation models and tools. Simulation has been used from 2012 onwards with the majority of publications between 2017 and 2020, being one of the first approaches tested to study the resilience of CPSs across multiple domains. To be mentioned is also the diffusion of its adoption, 84 articles in this review have used simulation-based analysis. Indeed simulation-based analyses are used also in other approaches, however, in this section simulations are the core of the research proposed and not only used to validate other methods and models. Bou-Harb et al. (2017) and Mouelhi (2019) utilize UPPAAL, a discrete event simulation tool. It is an integrated tool for modelling, simulation, and verification of real-time systems (Bengtsson, 1996). The objective of Bou-Harb et al. (2017) is to simulate the impact of a Distributed Denial of Service (DDoS) attack scenario on the CPS marine transportation. Their discrete event simulation model performed showing physical consequences, safety, and performance issues. Similarly, Mouelhi (2019) apply UPPAAL to analyse the resilience of CPS in a safety-critical scenario. Specifically, the tool helped the authors predicting endogenous resilience, defined as the *"Inherent ability of the system to detect and process internal faults and malicious attack"*. Their case study involves urban drone rescue systems: the simulation is used to predict and analyse the distributed interactive behaviour of the system, the timing constraints of its networking activities, and its survivability in case of functional and timing faults, or

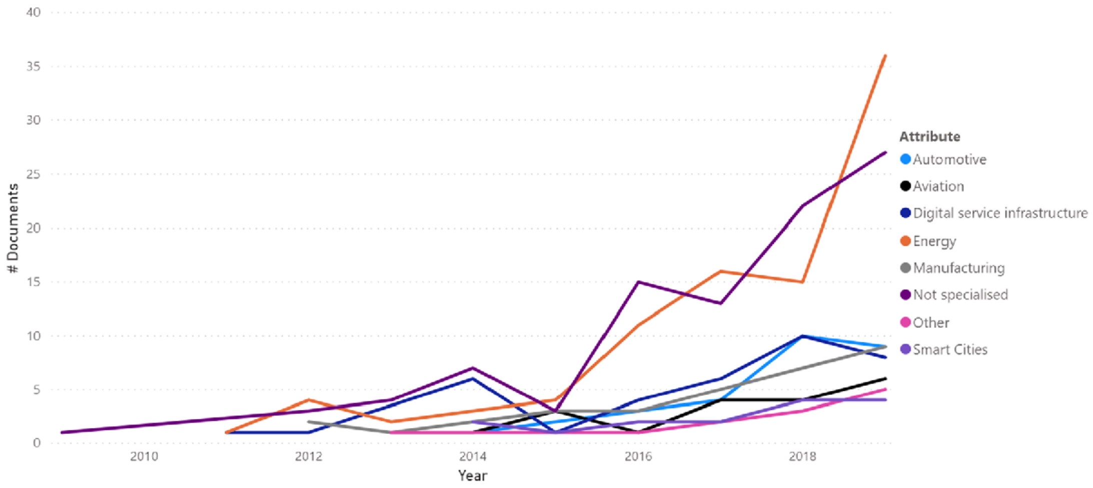


Fig. 5. Evolution of top domains of application.

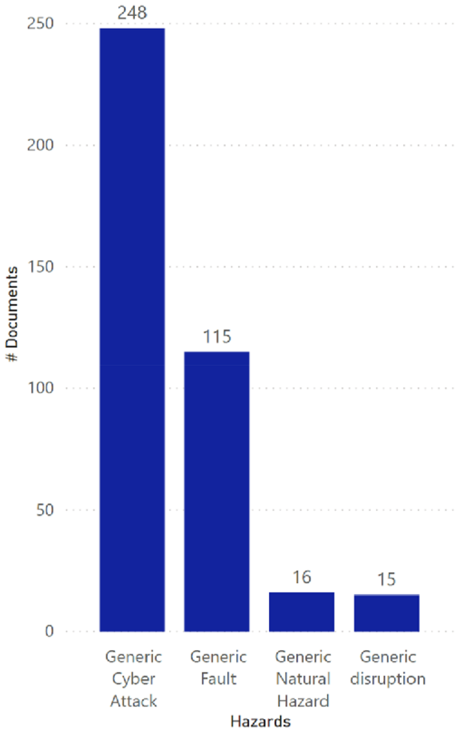


Fig. 6. Number of documents by hazards.

detected malicious attacks during operation.

Another time-discrete simulation software is developed by Egert (2017). The software is called HOLEG and it enables the simulation of Smart Grids (SGs) and examination of the state of SG over time. Moreover, HOLEG supports in controlling energy network resilience and helps in finding reconfiguration strategies for SGs through optimization algorithms. Similarly Koch, Moller and Deutschmann (2018) propose an event-based simulation via a Python-based simulation software that can predict the operability state of infrastructures under normal and disturbing conditions such as blackouts, power shortages, or cyber-attacks and therefore evaluate the impact on the critical infrastructure performance.

The reliable performance of the smart grid is addressed also in El Hariri (2019). The authors developed a hybrid hardware-software co-simulation platform to analyse the cyber and physical information flow in the smart grid and capture the interactions between the cyber and physical parts of the SG, The platform has been tested in three case studies: electrical vehicle charging control, power distribution network and finally a case study which tested a fake data injection attack. Co-simulation is tested also in an early-stage application in Fitzgerald, Pierce and Gamble (2012). Their focus is on co-simulation of discrete event models of controllers with continuous-time models of the controlled plant, joint with a model of the environment and the behaviours of interacting humans.

A different simulation model is suggested by Koutsoukos et al. (2018), Neema, Potteiger, Koutsoukos, Karsai, et al. (2018). They present the SecUre and REsilient Cyber-Physical Systems (SURE) platform which enables: models of cyber and physical components and their interactions, cyber-models that focus on the impact of attacks to CPS operation and operational scenarios useful for the evaluation of cyber-security risks. The platform has been validated in three case studies related to smart transportation systems.

Danilczyk, Sun and He (2019) and Yoginath (2019) instead present a Digital Twin for monitoring, surveillance and verification of two target CPS: the first involves a Canal Lock and the latter a Microgrid. A Digital Twin is a software-based replica of the real CPS and would perfectly monitor the current state of the system. In Yoginath (2019) the basis of the Digital Twin is Recurrent Neural Network (RNN)-based models. In Danilczyk et al. (2019) the Automatic Network Guardian for Electrical systems (ANGEL) Digital Twin using the legacy communication techniques found in the SCADA and wide-area monitoring systems can access real-time data and provide useful feedback.

Finally, a theoretical multi-agent approach is proposed in Kouicem, Raïevsky, and Occello (2020). In their work, which is still at an early stage, they suggest using a multi-agent simulation model to integrate knowledge from the psychology of emotion to improve cyber-physical systems resilience. Their model integrates artificial emotions to detect and anticipate abnormal situations, offers rapid recovery, and adapts the response behaviour of the situation.

4.2. *Game theory*

Traditional cybersecurity technologies are facing difficulties in managing complicated network traffic and they lack quantitative analysis and decision framework. In response to these considerations, interest in the application of game theory has been raised in recent years (Wang, 2017). Game theory shares many common issues with the cybersecurity problem. In game theory models, a player payoff depends

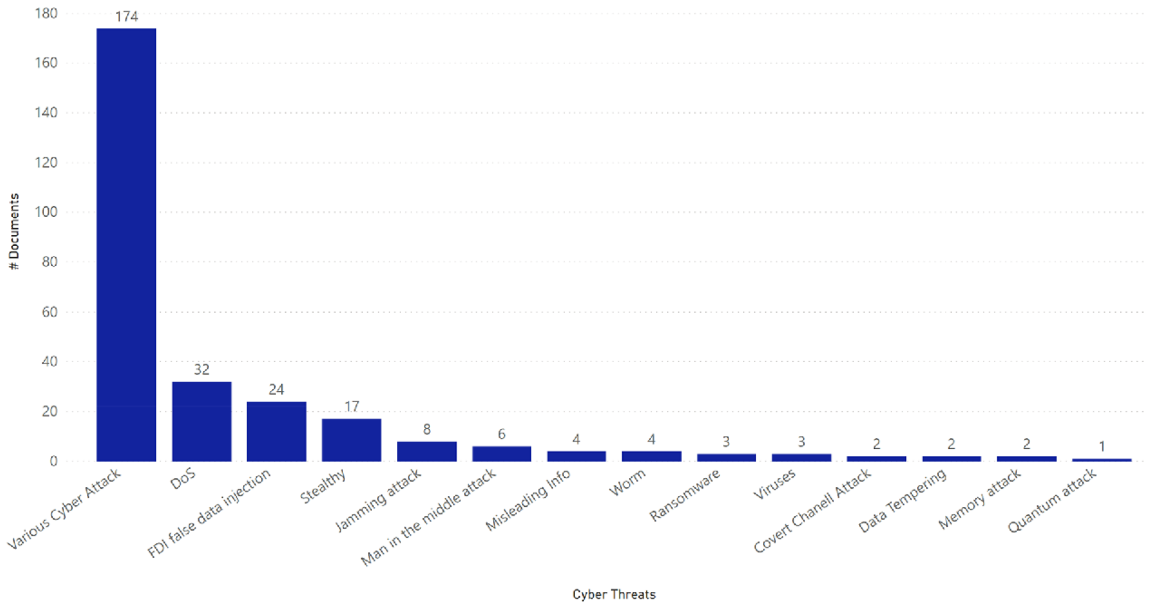


Fig. 7. Number of documents by cyber threats.

on both his own decisions and other players' behaviours as well as in cybersecurity problems depend on defence and attackers' strategies and network users' behaviours.

In this paragraph, different game-theoretic models and methods applied to the analysis of cyber-attack detection and mitigation actions with a focus on multiple domains are described.

Shen and Feng (2018) and Wu (2020) propose game-theoretic methods for robust and resilient control for CPSs under DoS attacks in industrial contexts. The resilient control problem for CPSs under DoS attack is studied also in Zhang (2019) in which a hierarchical game structure for the design of a resilient control system is proposed. An iterative zero-sum game is used in Zhu and Basar (2012) to model security policies at the cyber-level with corresponding optimal control response at the physical layer. A zero-sum, hybrid state stochastic game model is developed in Miao (2018) to detect and design defence mechanisms against multiple sensors attacks for CPSs. Yuan and Xia (2018) are also focusing on a resilient strategy for a class of CPS under DoS attack. Along with a minimax control strategy, a multi-channel transmission framework is established for transmitter and attacker and finally, defence strategies are obtained by solving a two-player Markov stochastic game.

Yuan, Sun, and Liu (2016) consider a resilient control system compromised by an intelligent DoS attacker, where a hierarchical Stackelberg game approach is used to integrate cybersecurity into the control system. On a different approach, Farraj (2016) concern the analysis of cyber switching attacks and control-based mitigation in smart grid systems. This paper integrates game theory with a smart grid dynamical system description to better study the system's behaviour in mitigating cyber-switching attacks. The resilience of the grid from cyber-physical attacks is also studied in the work of Baron-Prada, Osorio and Mojica-Nava (2017) a centralized transactive control algorithm for efficient coordination of microgrids based on population games theory. The solution they proposed was innovative from an analytical point of you, however, it resulted to be limited in the attack and microgrid' populations considered. Their algorithm was tested on a particular dynamic load altering attack and did not consider power storages such as houses, EV-cars Khalghani (2019).

4.3. *Graph theory*

Graph theory is commonly employed to model CPSs and critical infrastructures subjected to threats and disruption propagation (Inderwildi, 2020). With the network's nodes representing agents, components, and CPS's subsystems subjected to disruptions, and the network's edges representing the flow and intensity of disruption propagation. These approaches help explaining dependencies among nodes, modelling the supply and demand relationships of a network, and help in the decision-making phase after a disruption occurs. In this paragraph, different network models related to the resilience of cps are presented.

The complex coordination of infrastructures involved in power and water distribution networks has led to the spread of multiple network models that help in monitoring and mitigating attacks. Pasqualetti, Dörfler, and Bullo (2015) show examples both in the water distribution and power domain. Moreover, their study proposes multiple attack scenarios considering both the detection and analysis of the effects. Candelieri, Giordani and Archetti (2017) present a framework based on network analysis to enhance the resilience management in water distribution networks while also increasing the sustainability of CPS. In Hasan (2019) the authors propose a methodology to model the impact on safety and security energy delivery system (EDS). To do so, they develop a data-driven attack graph and fault-graph based model to evaluate the impact of threats in the CPS and criticality of the EDS nodes. Moreover, they propose an optimal resource allocation scheme of mitigation actions and policies. The results obtained in Wadhawan and Neuman (2016) through a graph-theoretic approach show how parameters such as time to criticality, power distribution capacity of the power grids nodes, and percentage of the nodes compromised help in defining the resilience of gas pipeline systems under cyber-physical attacks. Also, attack graphs are widely used to decide which set of security measures guarantee the safety of the system. Ibrahim and Alsheikh (2019) present a model-based attack graph for two communication networks which differ in their network topology. Wadhawan, Almajali, and Neuman (2018) instead develop a Bayesian Attack Graph for Smart Grid tool to compute the likelihood of the compromise of cyber components of smart grid. Leblanc and Koutsoukos (2018) and Fu (2020) investigate resilient consensus of first-order discrete-time. Fu (2020), first-order continuous-time, and higher-order continuous-time Leblanc and Koutsoukos (2018)

Table 1
Main cyber-attacks for CPSs emerged during literature review.

Cyber attack	Description	Reference
Denial-of-service (DoS) attack and Distributed-denial-of-service (DDoS) attack	DoS attacks aim at deteriorating the communication channels to prevent information exchange, usually either sensor data or control commands, between components of CPSs. In this type of attack, a huge volume of data is transmitted to the network to make the server busy, thereby disrupting normal services.	(Liu, Lu, & Wang, 2019; Ma, 2020; Sun, Zhang, & Shi, 2020)
Jamming attack	It is considered a kind of DoS attack and the most basic form of radio attack on a wireless communication system. It aims to disrupt a CPS preventing the control system from accessing current sensor measurements.	(Guan & Ge, 2018; Senejohnny, Tesi, & De Persis, 2018; Tomić, Breza, & McCann, 2019)
False data injection (FDI) attack	FDI attacks are considered among integrity attacks. In an FDI attack, an adversary could access and modify the physical system's state, sensor data, or control commands by introducing arbitrary errors and fake information. FDI attacks have been widely used against grid state estimation and energy management systems to disrupt the energy distribution.	(Guan & Ge, 2018; Ameli, 2020; Anubi, 2020)
Man-in-the-middle (MITM) attack	MITM attacks are part of the integrity attacks. Specifically, a MITM attack is an unauthorized interception and potential alteration of communication between two parties by an unauthorized third party. A representative attack of this kind in CPSs might disrupt data integrity establishing communication between the physical plant to the feedback controller through the network. Then, the original data is modified by the adversary sending false data to the feedback controller which processes it accordingly.	(Paridari, 2018; Jovanov & Pajic, 2019; Kim, 2019)
Stealthy attack	Stealthy attacks are sophisticated and potentially dangerous attacks. Usually conducted by skilled attackers who can penetrate control networks and then manipulate sensor readings or control signals persistently until the system crashes, while still keeping themselves undetected by following the expected behaviour of the system closely.	(Yang, 2018; Jovanov & Pajic, 2019; Zhou, 2019)
Malware attacks: ransomware, viruses, worm	A malware attack is a common cyberattack where malicious software can perform a variety of tasks and damage CPSs in multiple ways. It encompasses specific types of attacks such as ransomware, viruses, and worm. Ransomware denies or restricts access to victim files and demands a payment from the victim to restore access. Viruses are a type of malware that propagates by inserting a copy of itself and becoming part of another program. They spread from one computer to another, leaving infections as they travel. They can range in severity from causing mildly annoying effects to damaging data or software and causing denial-of-service (DoS) conditions. Unlike Viruses, Worms are spread via software vulnerabilities or phishing attacks. Worms and Viruses can modify and delete files, inject malicious software and steal data.	(Pajic et al., 2017; McDermott, 2019; Wedaj, Paul, & Ribeiro, 2019)

for complex cyber-physical networks subjected to attacks. Their idea is based on the concept that agents in the system influence one another by sharing information and for this reason a resilient consensus protocol - tested in the cases with and without trusted edges - is presented to achieve the resilient consensus and synchronization objectives.

4.4. System theoretic modelling

Among the most recent methods based on system-theory and able to understand safety, the System Theoretic Accident Model and Process (STAMP) is one of the most widely cited within academic literature (Allison, 2017). STAMP is a qualitative accident causation model created by Prof. Nancy Leveson to analyse accidents in systems (Leveson, 2004). Shintani, Aoyama, and Koshijima (2017) use a STAMP model to discuss the safety aspects of human in the loop CPS with humans as part of the CPS process. Their objective is to define a safe system that can guarantee human safety and a secure system able to respond to any threats. Additionally, the STAMP has been further extended by means of the System-Theoretic Process Analysis for Security (STPA-Sec). STPA-Sec is an extension of the systems safety-focused STPA, a STAMP-based hazard analysis developed for safety purposes, in the security domain. STPA-Sec defines the security problem at the system-level and uses control theory to design a feedback structure that identifies, reduces, and/or eliminates hostile system states (Leveson, 2017).

This approach is gaining increasing interest in recent literature. In Span et. al., (2018a), the authors offer a tailored version of the STPA-Sec approach for complex CPSs which objective is not limited to underline the security problem of assuring the system's critical function but also make the security problem more readily understandable to the stakeholder. This approach is tested in an exhaustive case study in Mailloux (2019), in which the tailored STPA-Sec approach is used for an autonomous space resupply vehicle. Span et al., (2018b) present an exhaustive case study of an STPA-Sec analysis for next-generation aerial refuelling systems. Moreover, the article offers a detailed description and recommendations of how to execute STPA-Sec for other complex systems.

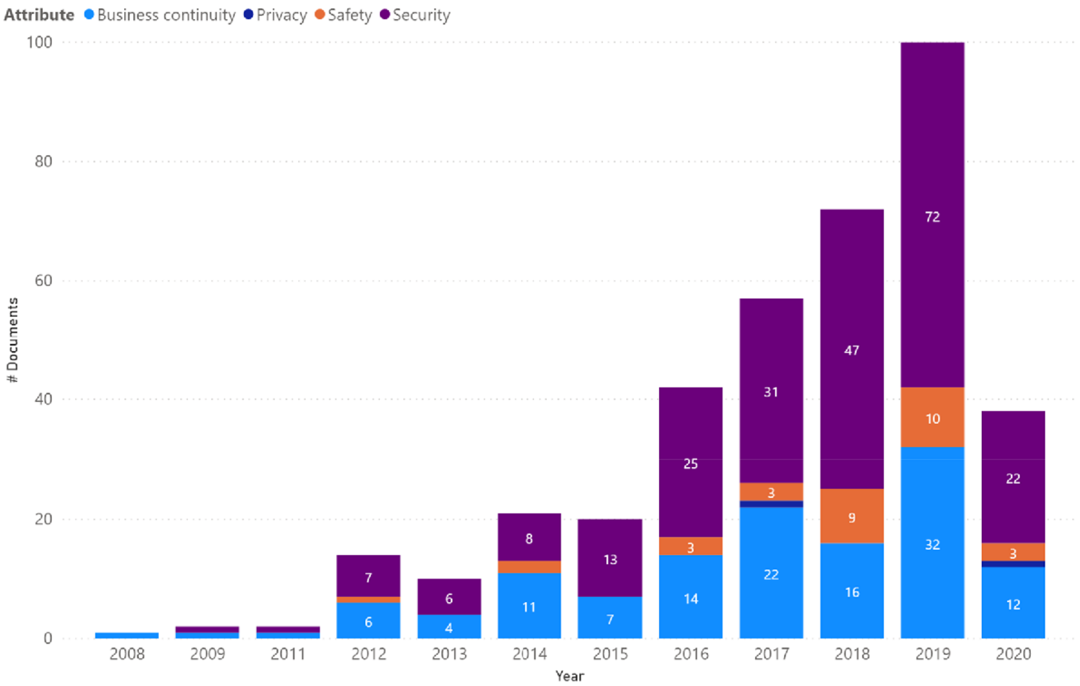


Fig. 8. Evolution of hazard effects.

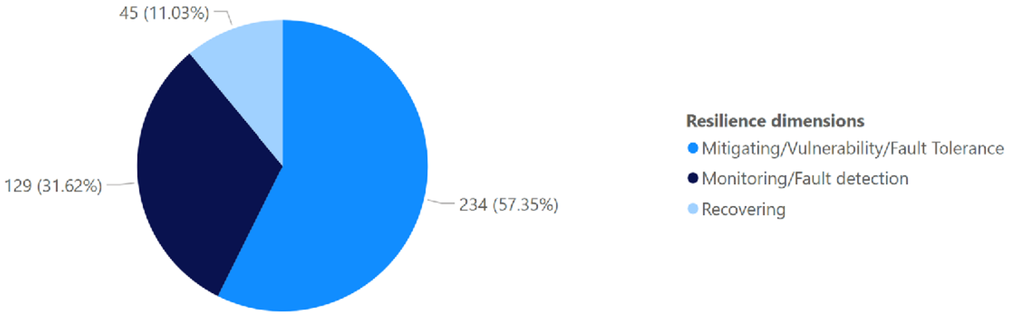


Fig. 9. Resilience dimensions.

Finally, Carter (2018) emphasize the need for solutions that study the security of cyber-physical systems as a safety problem. In CPS the *lines between the fields of safety and security become blurred* and in case of cyber-attacks, human lives involved in the safety-critical CPS might be in immediate danger. To meet this need, the authors propose a modified version of STPA-Sec which combines stakeholder perspectives and allows decision-makers to deal with safety, resilience, and security issues of CPS. Their solution is tested using a small UAV and a set of generated constraints.

4.5. Resilience metric

This section includes 24 contributions that deal with different metric-based approaches which are employed to evaluate resilience measures of individual properties of system components or functions. Metrics provide vital information of a system and are generally collected by analysing relevant attributes of that system (Linkov & Kott, 2018). The majority of these contributions are relatively recent, 18 publications are comprised between 2018 and 2020. This is probably due to the use of these metrics in complex systems involving multiple components and functions, e.g. smart grids (Almutairi, 2019; Venkataramanan et al., 2019; Venkataramanan, Hahn, & Srivastava, 2020), water treatment and distribution (Laszka, 2017; Murino, Armando, & Tacchella, 2019; Shin, 2020), intelligent transport systems (Halba, 2019; Laszka, 2020; Yang, 2020) which have undergone, as we have already mentioned in the bibliometric analysis, considerable growth in recent years. While first contributions proposing resilience measures focused on specific individual components, sensors, and CPS in general, not considering the system as a whole (Bai, Pasqualetti, & Gupta, 2015; Li & Kang, 2016; Park, Weimer, & Lee, 2017).

Murino et al. (2019), Laszka (2020) and Shin (2020) proposed different metrics to measure the resilience of water distribution networks and water treatment plants. In Murino et al. (2019) it is stressed the importance of a *"model-free, quantitative, and general-purpose evaluation methodology to extract resilience indexes"*. Their solution proposes a synthetic index that describes the amount of damage that a system can tolerate. The methodology is organized in three steps: identify the relevant state variables available from process control logs; build Figure Of Merit (FOM) functions considering (un)desirable values; compute resilience indexes based on FOM functions. Their approach does not detect or prevent a cyber-attack; however, they can underline significant deviation from the baseline and summarize the impact of an attack. Also, Shin (2020) address the vulnerability of the system, however, unlike Murino et al. (2019) which focus on a wastewater treatment plant, they study the resilience against cyber-physical attacks in water distribution systems. To do so, the authors propose a resilience metric that is a function of individual resilience capabilities. These can be identified in different system states before, during, and after disruptions. Specifically, in the proposed metric the following capabilities are considered: withstanding capability, absorptive capability, adaptive capability, and restorative capability. Also, Laszka (2020) reports a case study in the field of water distribution. However, its objective is to propose a metric that not only quantifies the impact of attacks but introduces a framework for prioritizing investment for reducing security risk. In their future work, they stress the difficulty of applying their framework to guide the design of robust power systems since these involve multiple components and are vulnerable to a large number of possible attacks which may be computationally expensive. The energy domain it is exactly what is discussed in the case study of Friedberg, McLaughlin and Smith (2017), Venkataramanan (2018), Clark and Zonouz (2019) and Venkataramanan, Hahn and Srivastava (2019, 2020). Clark and Zonouz (2019) proposes a resilience assessment metric for linear systems that focus on attacks that originate from the cyber network and then impact the physical components. They formulate a hierarchical game between the target CPS and a cyber threat, in which the value of each state of the game is equal to the resilience of the system. The metric uses competitive Markov decision processes to encode the dynamics and interdependencies of the cyber network along with the linear system models to capture the continuous dynamics of the physical processes. Venkataramanan et al. (2020) proposes a cyber-physical security assessment metric (CP-SAM) computed considering all the cyber-physical layers in the microgrid and the factors affecting its resilience, such as switching operations; redundant paths; probability of source; cyber vulnerabilities; status of device; network topology; quality of service, etc. The metric is then obtained by integrating these factors using fuzzy Choquet Integral. The Choquet integral is used also in Venkataramanan, Hahn and Srivastava (2019) in the decision-making process. The paper proposes a tool, named CyPhyR (cyber-physical resiliency) to help the operator and the planning engineer of a microgrid to improve its resilience. Two metrics are determined: The Cyber Asset Impact Potential (CAIP) metric in the planning phase and the Cyber Impact Severity (CIS) metric in the operational phase. CAIP helps in evaluating the criticality of each cyber asset used in the microgrid. CIS instead supports the operator in determining in real-time the resiliency of the microgrid. Friedberg et al. (2017) underlies the difficulty of designing effective resilience metrics which can take into consideration the diversity of challenges and performance measures. The authors identify seven requirements for a holist resilience metric and a descriptive resilience metric framework for smart grids. Finally Venkataramanan (2018) which demonstrate the application of cyber-physical resiliency metric able to quantify the impact on microgrid resilience. The metric is calculated with two components: a device level resiliency score based on the Trusted Safety Verifier and the Common Vulnerability Scoring Systems. The idea behind this metric is to uses vulnerability information available publicly and provide an intuitive score for the operator for monitoring the state of the system. Differently, Neema, Potteiger, Koutsoukos, Tang, et al. (2018) proposes a metrics-driven evaluation approach for evaluating the security and resilience of railways emphasizing the importance of implementing cybersecurity practices to ensure continuous safe operations in the presence of cyber-attacks. Also Whelihan (2017) focuses on transportation industry, specifically on unmanned aerial systems (UASs). The authors proposes a metric evaluation set as part of the Agile and Resilient Embedded Systems (ARES) methodology. ARES is a design-for-mission-assurance approach that focuses on system behaviours when essential functions are lost due to cyber-effect rather than focusing on the causes. Cybersecurity resilience requirements are pursuits through tangible and measurable innovative metrics. Their metric is grouped in a tuple metric consisting of detectability, isolability and recoverability characteristics. UASs are also considered, included in a wider network, in Sullivan, Colbert and Cowley (2018). The authors develop six cyber-resilience metrics for army tactical network architectures: time to detect degradation, percentage of functional services, lost contact time, percentage of successful guaranteed delivery messages, quality of voice and video over IP, time to switch to optimal modality. These previous examples offered domain-dependent system performance metrics to quantify resilience. In line with the cross-domain perspective of this research, Wang (2018) proposes a set of generic system performance metrics and a probabilistic design framework for CPS networks. These metrics are based on entropy and mutual information associated with the prediction and communication capabilities of networks in which each node corresponds to CPS components. When talking about networks, information, and communications channels it is worth mentioning the challenges involving the supply chain. Smart supply chains which involve CPSs are exposed to cyber threats. Chen, Dui, and Zhang (2020) attempt to derive an indicator based on the cost that helps in evaluating the supply chain resilience. The research affirms that the performance of the supply chain when interrupted is composed of three costs: cost of order loss, cost of order backlog, sales revenue, and cost of resilience ability. These are then combined in a cost-based supply chain resilience indicator which can be used to make decisions on the most suitable combination of resilience measures.

4.6. *Conceptual framework*

Multiple cyber-resilience frameworks have been proposed in the literature. The objective of this paragraph is to present and classify conceptual frameworks identified in the review. We will refer to conceptual framework as *"a network of linked concept that together provide a comprehensive understanding of the phenomena"* (Jabareen, 2009), that in this case refers to the CPSs resilience.

To better understand the purpose of these frameworks a possible classification in three levels is proposed, aligned with modern literature in resilience engineering. Levels are defined as micro, meso, and macro framework, in line with resilience literature (Bergström & Dekker, 2014). In this research, these three levels have been identified as follow: Micro-level research studies the smallest component such as sensors, IoT, and generically CPSs; Meso-level research studies frameworks that can be used for specific applications within the domains; Finally, macro-level research studies conceptual framework suitable for entire domains or critical infrastructures in general.

In Severson (2018) and Bin Masood (2019) two micro-level trust-based frameworks for closed-loop CPS are presented. Bin Masood (2019) addresses the potential of blockchain to implement a secure, trustworthy, and distributed management system for closed-loop CPS. Severson (2018) focuses the research on false data injection attacks that compromise multiple sensors. The trust-based consensus framework mitigates the impact of the attacks on a CPS. Moreover, this research stresses the relation between security and safety affirming that *"the most essential security priorities in a control system with distributed sensors are safety, integrity, and availability and develop their framework to reach these goals"*. Hahn (2015) state the lack of a common and unified attack analysis framework for CPS. Their framework is composed of a multi-layered reference architecture of the cyber, control, and physical layer of a CPS and a *meta*-model of cyber-attacks defined as cps kill-chain. This chain describes the possible cyber threats and the phases that make up an attack. Differently, Horowitz (2020), along with mechanisms of cyberattack detection, describes a framework for prioritizing alternative resilience solutions in CPS.

Moving to meso-level frameworks, Ashok, Govindarasu, and Wang (2017), Babiceanu and Seker (2017) and Naufal (2018) propose frameworks specific for an application. Ashok et al. (2017) focus on wide-area monitoring, protection, and control (WAMPAC) applications. The authors describe an end-to-end attack resilient security framework not only for risk assessment but useful also for attack prevention, detection mitigation, and resilience of the WAMPAC applications involved in the power grid. While this article focused on security as a cyber threat effect, the study of (Babiceanu & Seker, 2017) combines the theme of cybersecurity with the one of safety and business continuity. (Naufal, 2018) instead emphasizes how its conceptual framework is focused on increasing vehicle safety. Their idea is to develop an autonomous supervision and control system for intelligent transport systems (ITS). Moreover in their proposal, they underline the commonalities among CPSs, ITS and Roadway Transportation System (RTS) and the importance of a framework which includes multiple hazards and address suitable control actions to be taken to reduce the probability of accidents.

Interesting are also the macro-level conceptual frameworks of (Burns, 2018; Ahmadi-Assalemi, 2019) which focus on entire domains. A smart city is described as a system of systems (Burns, 2018) which includes complex systems such as smart buildings and smart workplaces. Smart business environments deal with the integration of sensors, robots, and humans which lead to the need for a Security-by-Design framework. In the research of Ahmadi-Assalemi (2019) it is stated that conventional access control systems and real-time detection are not sufficient to handle the latest cyber threats. Therefore, their framework incorporates blockchain technology which allows digital forensic readiness to facilitate post-incident investigations. Innovative is the use of CPSs objects as digital witnesses and logs generated by CPSs used for the process of even reconstruction. The pursuit of innovative solutions to strengthen the resilience of critical infrastructures is also carried out by (Baidya, Sun, & Perkins, 2019; Schneider, Dobie, & Ghettas, 2019). (Baidya et al., 2019) describe a conceptual framework based on social media and its sensors toward a resilient smart grid. Their idea is based on the fact that these tools can play a critical role during and after disasters for the role of sharing geotagged images, videos, textual feed but also information update and response planning. The research of (Schneider et al., 2019) instead, emphasizes the importance of leveraging approaches from other human-physical-operational risk systems such as process safety management (PSM) to manage the cybersecurity challenge of critical infrastructures. Finally, the work of (Ivanov, 2019) which focuses on cyber-physical supply chains and their related risks. Their solution departs from the supply chain risk management framework, dynamics, resilient, and control theoretic approaches. The derived framework shows that the integration of information technology – and therefore artificial intelligence, big data analytics, etc – into control theory can provide methods for dealing with the cyber supply chain.

4.7. *Algorithms*

This section proposes an overview of the different algorithms used to evaluate the resilience of CPSs. The first section focuses on resilient control algorithms and it is the richest in contributions. In this section the oldest contributions of the review are presented. The first work is dated 2008 and is the one of Woo (2008) which combines principles of feedback control and stability with methods for formal software engineering to design a system resilient to all classes of failures, both hardware and software. Instead, the first contribution to test a resilient control algorithm on a system subject to cyber-attack is that of Kottenstette, Karsai and Sztipanovits (2009). The authors are among the first to emphasize the importance of designing resilient control systems exploiting the intersection of computer science and control theory to mitigate new threats such as dos attacks. The theme of control systems and of the challenges related to new cyber threats, is treated extensively in the second paragraph on the design of attack resilient state estimators. Finally, a group of more recent contributions, from 2017 onwards that study machine learning techniques to measure the resilience of CPSs are presented.

4.7.1. *Resilient control algorithms*

Physical layers and computation elements of cyber-physical systems need a high level of coordination. This has led to an increase in computer interconnectivity over networks which is linked to both benefits and threats for critical infrastructures.

Hence, it became urgent to implement more resilient control algorithms that are both fault-tolerant and attack resilient (Sridhar, Hahn, & Govindarasu, 2012). Advanced control algorithms are dependent upon data collected from multiple sensors installed in resilient control systems; those data are further processed to monitor the state of the system and to maintain the stability of the system after fault occurrence. Among the first to talk about cps and resilient control algorithms is Abad, Caccamo and Robbins (2012) that in 2012 proposed a fault-resilient decentralized voltage control algorithm to make distributed cps, such as smart grids, resilient to communication faults. Also, Stibs (2018) focuses on cyber-physical energy systems. Their study proposes an iterative process for designing resilient distributed control algorithms for smart grids. The architecture described is based on four layers of defence mechanisms that help to increase the security of control algorithms. Even under the presence of malicious nodes, their system - inspired by blockchain theories - runs a decentralized control algorithm and needs to find a consensus. The case study tests the architecture simulating a Denial-Of-Service (DoS) attack. A different domain of application but the same hazard has been investigated by (Amullen, Shetty, & Keel, 2016). They developed a distributed model based on resilient control algorithms that enable each agent to accomplish their task even in the presence of a DoS attack that disrupt the communication among agents. In their case study, the agents considered to test the control algorithms are 6 Pioneer 3DX robots.

4.7.2. *Attack resilient state estimators*

The physical components of CPSs consist of parts such as controllers, actuators, and cyber components, such as sensors, which are the network that allows them to communicate. Due to this complex structure, CPSs are vulnerable to a range of cyber-attacks and thus require a robust defence and control system. However, most of the control systems have not been designed with cybersecurity in mind leading to a not reliable estimate of the true system states in case of an attack. Over the last decades, multiple incidents have involved control systems such as the cyber Stuxnet virus attack of an industrial supervisory control and data acquisition (SCADA) system of an Iran's nuclear plant in 2010 or the cyber-attack on a power control centre of Ukrainian electricity distribution network in 2015 stressing the importance of a resilient state estimation for control system (Zheng & Narasimha Reddy, 2017). The resilient state estimation problem is studied in (Fawzi, Tabuada, & Diggavi, 2014), considering a scenario in which some of sensors or actuators are attacked. In their estimation problem the authors give a characterization of the maximum number of attacks that can be detected and tolerated so that the state of the system can still be recovered. Their optimization problem resulted to be NP-hard. Also, in 2014, Pajic (2014) underlined the problem of state-estimators based on the *"non-realistic assumption that the exact system model is known"*. Their work can be considered an extension of (Fawzi et al., 2014) which computes a worst-case bound on the state estimate error in the presence of additive modelling errors with known bounds, but the optimization problem remains NP-hard. In this approach, the exact model of the system dynamics is not known, forcing the authors to simulate the solution on an unmanned grounded vehicle. In Pajic, Lee, and Pappas (2017) the same authors have further studied the problem in the case where there is bounded-size noise in the system's dynamics and the scenario represents the worst-case in which any signal can be injected via compromised sensors. (Marquis, 2018) takes into consideration the works of Pajic, Lee, and Pappas (2017) and proposes a resilient version of Kalman filtering and watermarking to not only estimate the correct state of the system but also to detect cyber-attacks. Differently from Pajic (2014) they assume that the system dynamics must be known beforehand and must be modelled as a linear time-invariant system. Their techniques resulted to be effective in the specific case of spoof and replay attacks. Yong, Zhu and Frazzoli (2015) instead propose an approach that does apply in the presence of unbounded noise signals and focuses on attacks on the network topology (e.g. switching attacks).

4.7.3. *Machine learning algorithms*

Other approaches identified to monitor the health of CPS and propose a real-time state estimation can be found in studies involving machine learning techniques. Amarasinghe (2018) propose a framework for measuring the operational health of a CPS and warn the operators of any degradation in cyber or physical health of the CPS. Their methodology starts with a process of data acquisition and feature extraction, the second step focuses on anomaly detection and possible state identification and state learning from historical data. Two unsupervised learning algorithms are used: Self-Organizing Maps and One-Class Support Vector Machines (OCSVM). Still considering strategies aimed at identifying anomalies in CPSs, Marino (2019) present a cyber-physical sensor called IREST (Industrial Control Systems Resilient Security Technology) which is able, using normal data through supervised (decision trees and random forest) and unsupervised (OCSVM) ML algorithms, to detect previously unseen disturbances.

Differently, Park, Weimer and Lee (2017) and Russell, Kwamena and Goubran (2019) use ML techniques respectively for sensor data validation and attacks on training data. Manipulation of sensors and data is a concern, trusting unconditionally data received might expose to cyber

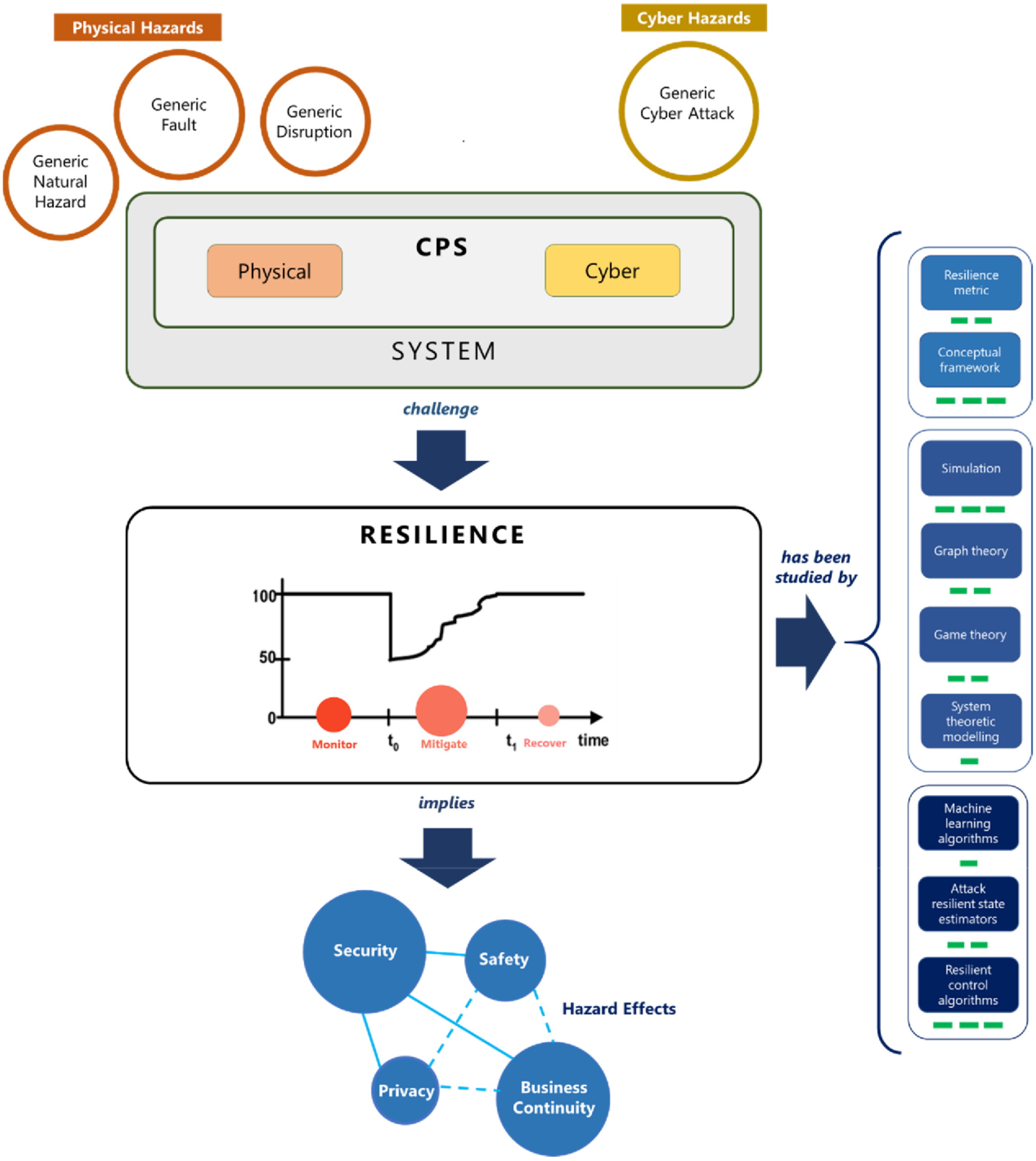


Fig. 10. Framework summarizing the findings of the survey. Larger bubble size implies higher number of contributions in the physical and cyber hazards (top); in the dimension of resilience being investigated (middle), and in the hazard effects (bottom). The rectangles on the right part of the figure define the prevalence of the listed approaches. Dotted connections among the hazard effects defines interactions with little to no exploration in current literature.

incidents. Authors in Russell et al. (2019) present fog-based sensor validation techniques using an artificial intelligence visual algorithm with already existing sensors to validate data and correct false predictions. Their method was tested on a smart building security scenario checking the robustness of cameras deployed to detect people. Park et al. (2017) state how data-driven CPS on one hand offers great capabilities for enhanced performance but on the other hand introduce vulnerability connected with the training data used for learning which can be tampered in several ways. With a case study in the healthcare domain, the authors propose a resilient classification algorithm trained under data attacks. Their results show that their 0–1 loss linear classification with a majority constraint is the most resilient algorithm among linear classification algorithms. Moreover, possible directions for countermeasures on training data attacks are addressed. Emphasis on the importance of assuring the safety of unmanned autonomous systems is instead addressed by Vachtsevanos (2018). The methodology proposed aims to detect and anticipate disturbances that may affect the operational integrity of complex systems. Their research works on detection and prediction of hazard evolution using prognostic and health management technologies, self-organization strategy for systems subjected to disturbances and reconfigurable control strategy using reinforcement learning techniques and finally a safety assurance and risk assessment methodology.

5. Discussion

The discussion starts from the bibliometric and descriptive findings to report the maturity of the domains, to identify threats, their evolution, and the effects that have been researched by scholars. Then, by developing an in-depth analysis of different methods, models, and frameworks, it is possible to understand their limitations and identify future research directions to deal with resilience of systems in which CPS are involved.

Fig. 10 summarizes the main aspects that emerged during the analysis and classification of the articles, in line with the previous contents. The size of the elements represents visually the frequency of threats, effects and resilience dimension researched in the articles and detailed in Section 3 and 4.

The elements in the upper part show how CPS are embedded in a larger, more complex system: the results of the analysis have underlined how CPSs can contribute to the resilience of technical and socio-technical systems only if they are well integrated with the context in which they operate. The analysis showed how each component in both physical facilities and cyber devices needs an appropriate level of security guard and it is challenged by multiple hazards. The review identified four categories of hazards, each represented by bubbles of different size, i.e. cyber-attacks, generic faults, generic natural hazards, generic disruption respectively standing for 248, 115, 16, 15 documents. These hazards challenge the system, which responds differently depending on its ability to withstand the threat. The majority of scholars have focused on developing solutions to enable systems to mitigate the effect of the threat, while others have instead focused on monitoring its state. This group includes mostly articles related to cyber-attacks and the development of tools able to detect an attack before it occurs. Finally, a limited number of studies suggest approaches for managing the post-attack phase and analyse the ability of CPSs to recover after the disruption occurs. It is interesting to note how the solutions available in literature to manage CPSs resilience are exploited in terms of different effects. A large number of articles, related to cyber hazards, as expected, focus on the topic of security and related investigation of CPS vulnerabilities and adversaries' characteristics. Only 4 of these articles have also highlighted issues related to privacy. The second largest group is the one related to business continuity, mainly related to physical hazards and only in recent studies extended towards cyber hazards. Finally, safety emerged as a dedicated subject over the last three years. In this context, it is worth noting that only recent contributions, as detailed in Section 4, are developing solutions to consider more than one effect at the same time.

Finally, the right part of the figure summarizes the answer to the second research question (i.e., eliciting available approaches to understand, measure, and model CPS resilience). Solutions relying on simulation models, and resilient control algorithms are mature and are reported in a large number of documents, while others, such as machine learning, are only emerging.

The next paragraphs detail the major themes emerged from literature, and discuss areas for potential future research in the area of resilience management in relation to CPSs.

- Section 5.1 details the diffusion of CPSs and highlights the domains that are currently under-researched.
- Section 5.2 discusses the benefit of including humans to improve the resilience of CPSs.
- Section 5.3 details the rationale for continuing to pursue the shift from cybersecurity to cyber-resilience and for investigating risks dynamically.
- Section 5.4 presents the major shortcomings of more recurring approaches in literature.
- Section 5.5 discusses the importance of a shift toward cyber-socio-technical systems by presenting the potential of STAMP model, and STPA-SEC technique.
- Section 5.6 extends the relevance of studying resilience in CPSs, as related to the recent pandemic outbreak

5.1. Diffusion of CPSs

What emerges clearly from the bibliometric analysis (cf. Fig. 3) is the increasing use and interest in CPSs over recent years. An awareness on the CPS relevance and utility is unbalanced across domains, for example energy research area shows maturity in terms of solutions able to measure the ability of these systems to detect, respond and recover after a disruption (cf. Fig. 4), while in the smart cities or transportation domain the analysis of resilience for CPS is rather emerging. An extensive discussion can be provided from a multi-domain perspective.

As extensively reported in Nguyen (2020), there is a variety of cyber detection and protection methods already in place and a rising interest in practices for structural resilience and operational resilience of smart energy systems. There is also a growing effort from academics towards the resilience of smart manufacturing systems. However, there is still a scarcity of literature and specifically, a lack of studies that focus on what to do after an intrusion has been detected to guarantee that these systems can continue operating safely.

The interest in smart transportation systems reports a growing trend from 2016 onwards. Initial studies are in the field of aviation. More recent ones focus on autonomous vehicles and intelligent transport systems. An expected dimension of investigation concerns the rising attention to safety issues concerning cyber threats (e.g.) (Naufal, 2018; Vachtsevanos, 2018; Van Wyk, 2020). Linked to the topic of intelligent transportation systems there is the one of smart cities. Smart cities, but also smart homes and smart workplaces are almost all reported in articles related to the last three years (2018–2020). Finally, it is interesting to underline that only two recent articles propose solutions to increase the resilience of smart supply chains. The state of the art on cyber risk in the supply chain, proposed by (Ghadge, 2019), underlines the lack of measures for cybersecurity in supply chains and the need for contextualized studies that address mitigation strategies for today's smart supply chains.

The diversity of domains being investigated suggests that future research should widen the scope of investigation towards hazards that go beyond security effects, but also includes safety, privacy, and business continuity.

5.2. The need for Human-in-the-loop in CPS

On this path, resilience analysis of large-scale socio-technical systems obviously includes humans, forcing research to move the unit of analysis from pure technical to socio-technical systems. Traditionally, the reviewed contributions identify human users as one of the major threats in today's technologies. Only few studies such as (Wu et al., 2019) or (Shintani et al., 2017) start considering humans to improve the effectiveness of CPSs resilience solutions. (Shintani et al., 2017) underlines the need of CPS to be structured in a way that ensures people to be incorporated as part of its process, defining Human-in-the-loop CPS. (Wu et al., 2019; Pinzone, 2020) instead, focusing on production systems, states that the objective of CPPSs should not consist of removing humans from the systems, but of involving them to take advantage of their intelligence. These studies constitute a promising field of research and future studies should start investigating more on quantitative models that consider *humans as part of the solution* (Zimmermann & Renaud, 2019).

5.3. Building resilience in cybersecurity

Other aspects to be investigated concern the threats modern systems face. The review confirms the expected dominance of cyber-attacks as threats for systems including CPSs. Nevertheless, these hazards are constantly changing, forcing a continuous reconsideration of strategies to ensure operational continuity. Referring to the framework proposed by the NIST (Institute of Standards, 2014), the reviewed articles show the increasing awareness of not limiting the analysis to the phases of detection and protection, but of considering as fundamental the ability to respond, recover and therefore withstand cyber incidents. Future studies should keep ensuring that the cyber defence mechanism is dynamic and real-time. Threats and incidents become more sophisticated, scholars need to keep pursuing the shift from cybersecurity to cyber-resilience understanding the nitty-gritty of these events in real operational settings, rather than simply fight them in small-world scenarios, managing variability rather than simply trying to reduce it.

Another aspect to consider is how to deal with these emerging threats in the discipline of risk assessment in the context of CPSs. Due to current turbulent operational environments, the sources of hazards are uncertain, and it is difficult to describe and model them quantitatively. However, at the same time, it is necessary to have quantitative measures for effective decision making and avoid that organizations are not able to react to new possible threats (Annarelli, Nonino, & Palombi, 2020). Current metrics frameworks usually fail at adapting to changes and, as already mentioned by (Zio, 2018), future studies should further investigate risk assessment dynamically, and moreover, integrate intentional cyber-attacks with natural hazards, human and organizational aspects.

5.4. System modelling and simulation

About quantitative approaches, there is still little knowledge about safety-related consequences of large-scale systems under cyber-attacks. In this context, future studies should encompass a greater use of system modelling and simulation, digital twin, graph theory, and system dynamics.

In particular, this review makes emerge that the use of Digital Twin is still not very diffuse, if compared to their widespread usage in other industrial settings. Future studies can take Digital Twin into consideration to monitor CPSs and evaluate possible intervention strategies. Same poor resonance among scholars is shared by System Dynamics, which on the contrary may constitute a valuable research method to evaluate the evolution of system performance in the case of an attack, even at different levels among the system being investigated.

5.5. The shift toward cyber-socio-technical systems

Moreover, for what concern quantitative resilience metrics, this review has shown the proliferation of indicators that are inherently domain specific. On the other hand, resilience metrics that consider both technical and socio-technical aspects are limitedly discussed in literature. Therefore, it emerges the need for a framework of flexible socio-technical metrics to evaluate the resilience of systems including CPSs (Patriarca, 2021). While there is a general agreement that resilience is something a system does (rather than something a system has), it is inherently system specific. However, at certain abstraction levels, metrics can be remaining coherent to allow a general understanding based on dimensions that could correspond to different resilience abilities (e.g., monitoring, absorbing, recovery).

Finally, a dedicated mention can be provided about the usage of STAMP and its hazard analysis technique, i.e., STPA-SEC, in the pursuit of resilience for systems with CPSs. STAMP is a valid and widely established model in safety literature, however not yet widespread in cyber-resilience. Considering the increased interest, described above, in the safety of CPSs, applications of STAMP and STPA-SEC can help at integrating safety and security needs. This is particularly interesting since STAMP differs from other cybersecurity approaches in the way it maps interdependencies between system elements, both technical and human. As such, it represents a promising research area for future socio-technical cyber analyses.

5.6. CPSs response to pandemic outbreak

The rapid and worldwide expansion of the coronavirus has forced us to revisit our understanding of the resilience of various systems, CPSs make no exception. IT has played a positive central role in all activities, serving as the focal point of operations in healthcare, business, education, industry, and more (Weil & Murugesan, 2020). However, there are several drawbacks, such as increased cybersecurity threats and hazards, performance problems as a result of a massive increase in workload, and business continuity. The pandemic highlighted the vulnerability of traditional cybersecurity systems (Ramadan, 2021). Limitations and weaknesses concerning CPS resilience discussed in this review have become more visible. Dealing with a highly dynamic and uncertain environment demands for approaches that emphasize resilience and channel risk mitigation efforts across a broad boundary, contemplating CPSs, autonomous systems, and remote workers (Schoitsch, 2020). First, the resilience of digital infrastructure must be assessed from both a social and technical lens, considering systems from a socio-technical security point of view (Gardner-Stephen & Nabben, 2020). This observation does not only refer to secure endpoints used by smart workers, but also promotes prevention mechanisms involving people, processes, and procedures to react to social engineering attacks (Ramadan, 2021). Second, it is necessary to ensure that incident response protocols reflect the new operating conditions.

Another aspect on which there has been much debate during the pandemic is business continuity. Three scenarios have been documented: a few organizations had excellent business continuity plans, some did not have them at all, and many had plans but did not consider pandemic among the risk scenarios (Galbusera, Cardarilli, & Giannopoulos, 2021). Among the most significant shortcomings observed in analyzing business continuity plans are: lack of documentation, formal process mapping and backups; usage of individual applications not integrated with the company's information system; reduced responsiveness in equipping infrastructure for new traffic and usage patterns (Ivanov & Dolgui, 2020; Papagiannidis, Harris, & Morton, 2020; Galbusera et al., 2021). With regard to the latter point, several studies emphasize that organizations relying on cloud-based services and infrastructures even before the pandemic were found to be the quickest to readjust, to scale solutions such as bandwidth and computing capacity, and easily maintain continuity of operations (Umar, 2020; Zhang, 2021). Finally, going to evaluate the upsides, this situation has led to a general increase in the awareness that CPSs, IoT networks and integrated intelligence are considered the drivers of innovation that need to be responsibly incorporated into industrial contexts and social life to enable a resilient and sustainable society. Some examples are provided in the work by (Gupta et al., 2021) who highlight how community resilience to pandemic can only be achieved through integrated IoTs and CPSs architecture implemented in smart communities (e.g., hospitals, homes, transportation etc.). Other considerations are included in the work by Weil and Murugesan (2020) and Wuest (2020) who use examples coming from factories that already embraced a digital transformation to emphasize the technologies that allowed them to be more responsive. Finally, the research by Li (2020) proposes more details about an intelligent manufacturing systems framework to address the challenges and potential for existing manufacturing technologies to fight similar disruptions and underline the importance of integration between technologies, business systems, and humans.

6. Conclusion

CPSs empower modern systems helping the diffusion of smart services. CPSs give us new opportunities, functionalities and facilitate disaster response, however they have also enabled new forms of threats and expose systems to new vulnerabilities.

The use of CPSs and their integration with physical systems is fundamental and as such it requires a thorough investigation. In the past, scholars tried to address the problem through risks identification and protection, and detection. Today it is definite the need to move towards a resilience prospective.

This review has confirmed the benefits as well as the potential criticality of CPSs which need to be designed in a way that ensures system resilience. The mapped evidence from this survey highlights the great prevalence of research that investigates CPSs resilience at a technical level while new emerging domains require a joint technical and socio-technical research dimension which highlights promising research opportunities. We believe that a cross-domain multi-method perspective may foster scholars achieving this target.

Acknowledgement

This work is supported by the fund “Progetto di Eccellenza” of the Department of Computer, Control and Management Engineering “Antonio Ruberti” of Sapienza University of Rome. The department has been designated by the Italian Ministry of Education (MIUR) for being “Department of Excellence” in advanced training programs in the field of cybersecurity.

References

Abad, F. A. T., Caccamo, M., & Robbins, B. (2012). A Fault Resilient Architecture for Distributed Cyber-Physical Systems. In *2012 IEEE International Conference on Embedded and Real-Time Computing Systems and Applications. 2012 IEEE 18th International Conference on Embedded and Real-Time Computing Systems and Applications (RTCSA*. https://doi.org/10.1109/rtcsa.2012.18

Abraham, J., Banerjee, S., & Chatterjee, A. (2017). Design of efficient error resilience in signal processing and control systems: From algorithms to circuits. In *2017 IEEE 23rd International Symposium on On-Line Testing and Robust System Design (IOLTS). 2017 IEEE 23rd International Symposium on On-Line Testing and Robust System Design (IOLTS)*. https://doi.org/10.1109/iolts.2017.8046241

Ahmadi-Assalemi, G., al-Khateeb, H. M., Epiphaniou, G., Cosson, J., Jahankhani, H., & Pillai, P. (2019, January). Federated Blockchain-Based Tracking and Liability Attribution Framework for Employees and Cyber-Physical Objects in a Smart Workplace. In 2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3). 2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3). Doi: 10.1109/icgs3.2019.8688297.

Allison, C. K., Revell, K. M., Sears, R., & Stanton, N. A. (2017). Systems Theoretic Accident Model and Process (STAMP) safety modelling applied to an aircraft rapid decompression event. *Safety Science, 98*, 159–166. https://doi.org/10.1016/j.ssci.2017.06.011

Almutairi, A., Wheeler, J. P., Slutzky, D. L., & Lambert, J. H. (2019). Integrating Stakeholder Mapping and Risk Scenarios to Improve Resilience of Cyber-Physical-Social Networks. *Risk Analysis, 39*(9), 2093–2112. https://doi.org/10.1111/risa.13292

Amarasinghe, K., Wickramasinghe, C., Marino, D., Rieger, C., & Manic, M. (2018). August). Framework for Data Driven Health Monitoring of Cyber-Physical Systems. In *2018 Resilience Week (RWS). 2018 Resilience Week (RWS)*. https://doi.org/10.1109/rweek.2018.8473535

Ameli, A., Hooshyar, A., El-Saadany, E. F., & Youssef, A. M. (2020). An Intrusion Detection Method for Line Current Differential Relays. *IEEE Transactions on Information Forensics and Security, 15*, 329–344. https://doi.org/10.1109/tifs.2019.2916331

Amullen, E. M., Shetty, S., & Keel, L. H. (2016, July). Model-based resilient control for a multi-agent system against Denial of Service attacks. 2016 World Automation Congress (WAC). In 2016 World Automation Congress (WAC). Doi: 10.1109/wac.2016.7582963.

Annarelli, A., Nonino, F., & Palombi, G. (2020). Understanding the management of cyber resilient systems. *Computers and Industrial Engineering, 149*. https://doi.org/10.1016/j.cie.2020.106829

Anubi, O. M., Konstantinou, C., Wong, C. A., & Vedula, S. (2020). August). Multi-Model Resilient Observer under False Data Injection Attacks. In *2020 IEEE Conference on Control Technology and Applications (CCTA). 2020 IEEE Conference on Control Technology and Applications (CCTA)*. https://doi.org/10.1109/ccta41146.2020.9206284

Arghandeh, R., von Meier, A., Mehrmanesh, L., & Mili, L. (2016). On the definition of cyber-physical resilience in power systems. *Renewable and Sustainable Energy Reviews, 58*, 1060–1069. https://doi.org/10.1016/j.rser.2015.12.193

Ashok, A., Govindarasu, M., & Wang, J. (2017). Cyber-Physical Attack-Resilient Wide-Area Monitoring, Protection, and Control for the Power Grid. In *Proceedings of the IEEE*. https://doi.org/10.1109/JPROC.2017.2686394

Babiceanu, R. F., & Seker, R. (2017). Trustworthiness Requirements for Manufacturing Cyber-physical Systems. *Procedia Manufacturing*. https://doi.org/10.1016/j.promfg.2017.07.202

Bai, C.-Z., Pasqualetti, F., & Gupta, V. (2015). Security in stochastic control systems: Fundamental limitations and performance bounds. In *2015 American Control Conference (ACC). 2015 American Control Conference (ACC)*. https://doi.org/10.1109/acc.2015.7170734

Baidya, P. M., Wei Sun, & Perkins, A. (2019). A Survey on Social Media to Enhance the Cyber-PhysicalSocial Resilience of Smart Grid. 8th Renewable Power Generation Conference (RPG 2019). 8th Renewable Power Generation Conference (RPG 2019). Doi: 10.1049/cp.2019.0602.

Baron-Prada, E., Osorio, E., & Mojica-Nava, E. (2017). October). Resilient transactive control in microgrids under dynamic load altering attacks. In *2017 IEEE 3rd Colombian Conference on Automatic Control (CCAC). 2017 IEEE 3rd Colombian Conference on Automatic Control (CCAC)*. https://doi.org/10.1109/ccac.2017.8276400

Bengtsson, J., Larsen, K., Larsson, F., Pettersson, P., & Yi, W. (1996). UPPAAL —a tool suite for automatic verification of real-time systems. In Hybrid Systems III (pp. 232–243). Springer Berlin Heidelberg. Doi: 10.1007/bfb0020949.

Bennaceur, A., Ghezzi, C., Tei, K., Kehrer, T., Weyns, D., Calinescu, R., Dustdar, S., Hu,Z., Honiden, S., Ishikawa, F., Jin, Z., Kramer, J., Litoiu, M., Loreti, M., Moreno, G., Muller, H., Nenzi, L., Nuseibeh, B., Pasquale, L., … Zhao, H. (2019, May). Modelling and Analysing Resilient Cyber-Physical Systems. In 2019 IEEE/ACM 14th International Symposium on Software Engineering for Adaptive and Self-Managing Systems (SEAMS). Doi: 10.1109/seams.2019.00018.

Bergström, J., Dekker, S.W.A., 2014. Bridging the macro and the micro by considering the meso: Reflections on the fractal nature of resilience. Ecology and Society. Resilience Alliance 19 (4)Doi: 10.5751/ES-06956-190422.

Bin Masood, A., Qureshi, H. K., Danish, S. M., & Lestas, M. (2019). Realizing an Implementation Platform for Closed Loop Cyber-Physical Systems Using Blockchain. In *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring). 2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*. https://doi.org/10.1109/vtcspring.2019.8746372

pp. Björck, F., Henkel, M., Stirna, J., & Zdravkovic, J. (2015). Cyber Resilience – Fundamentals for a Definition. In New Contributions in Information Systems and Technologies (pp. 311–316). Springer International Publishing. Doi: 10.1007/978-3-319-16486-1_31.

Bou-Harb, E., Kaisar, E. I., & Austin, M. (2017, June). On the impact of empirical attack models targeting marine transportation. 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS). 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS). Doi: 10.1109/mtits.2017.8005665.

Boyes, H. A. (2013). Trustworthy cyber-physical systems - a review. 8th IET International System Safety Conference incorporating the Cyber Security Conference 2013. Doi: 10.1049/cp.2013.1707.

Burns, M., Griffor, E., Balduccini, M., Vishik, C., Huth, M., & Wollman, D. (2018). June). Reasoning about Smart City. In *2018 IEEE International Conference on Smart Computing (SMARTCOMP). 2018 IEEE International Conference on Smart Computing (SMARTCOMP)*. https://doi.org/10.1109/smartcomp.2018.00033

Candelieri, A., Giordani, I., & Archetti, F. (2017). Supporting Resilience Management of Water Distribution Networks through Network Analysis and Hydraulic Simulation. In *2017 21st International Conference on Control Systems and Computer Science (CSCS). 2017 21st International Conference on Control Systems and Computer Science (CSCS)*. https://doi.org/10.1109/cscs.2017.91

Carter, B. T., Bakirtzis, G., Elks, C. R., & Fleming, C. H. (2018). A systems approach for eliciting mission-centric security requirements. In *2018 Annual IEEE International Systems Conference (SysCon). 2018 Annual IEEE International Systems Conference (SysCon)*. https://doi.org/10.1109/syscon.2018.8369539

Chen, L., Dui, H., & Zhang, C. (2020). A resilience measure for supply chain systems considering the interruption with the cyber-physical systems. *Reliability Engineering and System Safety, 199*. https://doi.org/10.1016/j.ress.2020.106869

pp. Cheng, P., Cui, A., Yang, Y., Luo, Y., & Sun, W. (2017, October). Recognition and classification of coating film defects on automobile body based on image processing. 2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). Doi: 10.1109/cisp-bmei.2017.8302070.

Clark, A., & Zonouz, S. (2019). Cyber-physical resilience: Definition and assessment metric. *IEEE Transactions on Smart Grid*. https://doi.org/10.1109/TSG.2017.2776279

Danilczyk, W., Sun, Y., & He, H. (2019). ANGEL: An Intelligent Digital Twin Framework for Microgrid Security. In *2019 North American Power Symposium (NAPS). 2019 North American Power Symposium (NAPS)*. https://doi.org/10.1109/naps46351.2019.9000371

Das, L., Munikoti, S., Natarajan, B., & Srinivasan, B. (2020). Measuring smart grid resilience: Methods, challenges and opportunities. *Renewable and Sustainable Energy Reviews, 130*, Article 109918. https://doi.org/10.1016/j.rser.2020.109918

Dibaji, S. M., Pirani, M., Flamholz, D. B., Annaswamy, A. M., Johansson, K. H., & Chakrabortty, A. (2019). A systems and control perspective of CPS security. *Annual Reviews in Control, 47*, 394–411. https://doi.org/10.1016/j.arcontrol.2019.04.011

Ding, D., Han, Q.-L., Xiang, Y., Ge, X., & Zhang, X.-M. (2018). A survey on security control and attack detection for industrial cyber-physical systems. *Neurocomputing, 275*, 1674–1683. https://doi.org/10.1016/j.neucom.2017.10.009

Egert, R., Cordero, C. G., Tundis, A., & Muhlhauser, M. (2017). October). HOLEG: A simulator for evaluating resilient energy networks based on the Holon analogy. In *2017 IEEE/ACM 21st International Symposium on Distributed Simulation and Real-Time Applications (DS-RT)*. https://doi.org/10.1109/distra.2017.8167665

El Hariri, M., Youssef, T., Saleh, M., Faddel, S., Habib, H., & Mohammed, O. A. (2019). A Framework for Analyzing and Testing Cyber-Physical Interactions for Smart Grid Applications. *Electronics, 8*(12), 1455. https://doi.org/10.3390/electronics8121455

Farraj, A., Hammad, E., Daoud, A. A., & Kundur, D. (2016). A Game-Theoretic Analysis of Cyber Switching Attacks and Mitigation in Smart Grid Systems. *IEEE Transactions on Smart Grid, 7*(4), 1846–1855. https://doi.org/10.1109/tsg.2015.2440095

Fawzi, H., Tabuada, P., & Diggavi, S. (2014). Secure estimation and control for cyber-physical systems under adversarial attacks. *IEEE Transactions on Automatic Control*. https://doi.org/10.1109/TAC.2014.2303233

Fitzgerald, J., Pierce, K., & Gamble, C. (2012, June). A rigorous approach to the design of resilient cyber-physical systems through co-simulation. IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN 2012). 2012 IEEE/IFIP 42nd International Conference on Dependable Systems and Networks Workshops (DSN-W).Doi: 10.1109/dsnw.2012.6264663.

Fraga-Lamas, P., & Fernández-Caramés, T. M. (2019). A Review on Blockchain Technologies for an Advanced and Cyber-Resilient Automotive Industry. *IEEE Access*. https://doi.org/10.1109/ACCESS.2019.2895302

Friedberg, I., McLaughlin, K., & Smith, P. (2017). April). A cyber-physical resilience metric for smart grids. In *2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT). 2017 IEEE Power & Energy Society Innovative Smart Grid Technologies Conference (ISGT)*. https://doi.org/10.1109/isgt.2017.8086065

Fu, W., Qin, J., Shi, Y., Zheng, W. X., & Kang, Y. (2020). Resilient Consensus of Discrete-Time Complex Cyber-Physical Networks Under Deception Attacks. *IEEE Transactions on Industrial Informatics, 16*(7), 4868–4877. https://doi.org/10.1109/tii.2019.2933596

Galbusera, L., Cardarilli, M., Giannopoulos, G., 2021. The ERNCIP survey on COVID-19: Emergency & Business Continuity for fostering resilience in critical infrastructures. Safety Science Doi: 10.1016/j.ssci.2021.105161.

Gardner-Stephen, P., Nabben, K., 2020. Capacity Maintenance during Global Disruptions: Security, resilience and incentives matter. In: 2020 IEEE Global Humanitarian Technology Conference, GHTC 2020. Doi: 10.1109/GHTC46280.2020.9342865.

Ghadge, A., Weiß, M., Caldwell, N. D., & Wilding, R. (2019). Managing cyber risk in supply chains: A review and research agenda. *Supply Chain Management: An International Journal, 25*(2), 223–240. https://doi.org/10.1108/scm-10-2018-0357

Guan, Y., & Ge, X. (2018). Distributed Attack Detection and Secure Estimation of Networked Cyber-Physical Systems Against False Data Injection Attacks and Jamming Attacks. *IEEE Transactions on Signal and Information Processing over Networks*. https://doi.org/10.1109/TSIPN.2017.2749959

Gujrati, S., Zhu, H., & Singh, G. (2015). In *August). Composable Algorithms for Interdependent Cyber Physical Systems*. https://doi.org/10.1109/rweek.2015.7287431

Gupta, D., Bhatt, S., Gupta, M., & Tosun, A. S. (2021). Future Smart Connected Communities to Fight COVID-19 Outbreak. *Internet of Things, 13*, Article 100342. https://doi.org/10.1016/j.iot.2020.100342

Gupta, R., Tanwar, S., Al-Turjman, F., Italiya, P., Nauman, A., & Kim, S. W. (2020). Smart Contract Privacy Protection Using AI in Cyber-Physical Systems: Tools, Techniques and Challenges. *IEEE Access, 8*, 24746–24772. https://doi.org/10.1109/access.2020.2970576

Haggi, H., nejad, R. R., Song, M., & Sun, W. (2019, May). A Review of Smart Grid Restoration to Enhance Cyber-Physical System Resilience. 2019 IEEE Innovative Smart Grid Technologies - Asia (ISGT Asia). 2019 IEEE Innovative Smart Grid Technologies – Asia (ISGT Asia). Doi: 10.1109/isgt-asia.2019.8881730.

Hahn, A., Thomas, R. K., Lozano, I., & Cardenas, A. (2015). A multi-layered and kill-chain based security analysis framework for cyber-physical systems. *International Journal of Critical Infrastructure Protection, 11*, 39–50. https://doi.org/10.1016/j.ijcip.2015.08.003

Halba, K., Griffor, E., Kamongi, P., & Roth, T. (2019). October). Using Statistical Methods and Co-Simulation to Evaluate ADS-Equipped Vehicle Trustworthiness. In *2019 Electric Vehicles International Conference (EV). 2019 Electric Vehicles International Conference(EV)*. https://doi.org/10.1109/ev.2019.8892870

Hasan, K., Shetty, S., Hassanzadeh, A., & Ullah, S. (2019). November). Towards Optimal Cyber Defense Remediation in Cyber Physical Systems by Balancing Operational Resilience and Strategic Risk. In *MILCOM 2019–2019 IEEE Military Communications Conference (MILCOM)*. https://doi.org/10.1109/milcom47813.2019.9021076

Horowitz, B. M. (2020). Cyberattack-Resilient Cyberphysical Systems. *IEEE Security & Privacy, 18*(1), 55–60. https://doi.org/10.1109/msec.2019.2947123

Ibrahim, M., & Alsheikh, A. (2018). Assessing Level of Resilience Using Attack Graphs. In *2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI). 2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*. https://doi.org/10.1109/ecai.2018.8679044

Inderwildi, O., Zhang, C., Wang, X., & Kraft, M. (2020). The impact of intelligent cyber-physical systems on the decarbonization of energy. *Energy & Environmental Science, 13*(3), 744–771. https://doi.org/10.1039/c9ee01919g

Institute of Standards, N.. (2014). 'Framework for Improving Critical Infrastructure Cybersecurity. *Version, 1.1'*. https://doi.org/10.6028/NIST.CSWP.04162018

Ivanov, D., Dolgui, A., Sokolov, B., & Ivanova, M. (2019). Intellectualization of control: Cyber-physical supply chain risk analytics. *IFAC-PapersOnLine, 52*(13), 355–360. https://doi.org/10.1016/j.ifacol.2019.11.146

Ivanov, D., Dolgui, A., 2020. A digital supply chain twin for managing the disruption risks and resilience in the era of Industry 4.0. Production Planning and Control Doi: 10.1080/09537287.2020.1768450.

Jabareen, Y. (2009). Building a Conceptual Framework: Philosophy, Definitions, and Procedure. *International Journal of Qualitative Methods, 8*(4), 49–62. https://doi.org/10.1177/160940690900800406

Jacobs, N., Hossain-McKenzie, S., & Vugrin, E. (2018). In *August). Measurement and Analysis of Cyber Resilience for Control Systems: An Illustrative Example*. https://doi.org/10.1109/rweek.2018.8473549

Jovanov, I., & Pajic, M. (2019). Relaxing Integrity Requirements for Attack-Resilient Cyber-Physical Systems. *IEEE Transactions on Automatic Control*. https://doi.org/10.1109/TAC.2019.2898510

Khalghani, M. R., Solanki, J., Solanki, S. K., & Sargolzaei, A. (2019). August). Resilient and Stochastic Load Frequency Control of Microgrids. In *2019 IEEE Power & Energy Society General Meeting (PESGM). 2019 IEEE Power & Energy Society General Meeting (PESGM)*. https://doi.org/10.1109/pesgm40551.2019.8974111

Khalid, F., Rehman, S., & Shafique, M. (2020). Overview of Security for Smart Cyber-Physical Systems. In *Security of Cyber-Physical Systems* (pp. 5–24). Springer International Publishing. https://doi.org/10.1007/978-3-030-45541-5_2.

Kim, S., Won, Y., Park, I.-H., Eun, Y., & Park, K.-J. (2019). Cyber-Physical Vulnerability Analysis of Communication-Based Train Control. *IEEE Internet of Things Journal, 6*(4), 6353–6362. https://doi.org/10.1109/jiot.2019.2919066

Koch, T., Moller, D. P. F., & Deutschmann, A. (2018). A Python-Based Simulation Software for Monitoring the Operability State of Critical Infrastructures Under Emergency Conditions. In *2018 IEEE International Conference on Electro/Information Technology (EIT)*. https://doi.org/10.1109/eit.2018.8500219

Kottenstette, N., Karsai, G., & Sztipanovits, J. (2009, August). A passivity-based framework for resilient cyber physical systems. 2009 2nd International Symposium on Resillient Control Systems. 2009 2nd International Symposium on Resillient Control Systems (ISRCS). Doi: 10.1109/isrcs.2009.5251370.

Kouicem, E., Raïevsky, C., & Occello, M. (2020). Towards a cyber-physical systems resilience approach based on artificial emotions and multi-agent systems. *Paper presented at the ICAART 2020 - Proceedings of the 12th International Conference on Agents and Artificial Intelligence*. Retrieved from www.scopus.com.

Koutsoukos, X., Karsai, G., Laszka, A., Neema, H., Potteiger, B., Volgyesi, P., et al. (2018). SURE: A Modeling and Simulation Integration Platform for Evaluation of Secure and Resilient Cyber-Physical Systems. *Proceedings of the IEEE, 106*(1), 93–112. https://doi.org/10.1109/jproc.2017.2731741

Laszka, A., Abbas, W., Vorobeychik, Y., & Koutsoukos, X. (2017, April 21). Synergic security for smart water networks. Proceedings of the 3rd International Workshop on Cyber- Physical Systems for Smart Water Networks. CPS Week '17: Cyber Physical Systems Week 2017. Doi: 10.1145/3055366.3055376.

Laszka, A., Abbas, W., Vorobeychik, Y., & Koutsoukos, X. (2019). *Integrating redundancy, diversity, and hardening to improve security of industrial internet of things.Cyber-Physical Systems, 6*(1), 1–32. https://doi.org/10.1080/23335777.2019.1624620

Leblanc, H. J., & Koutsoukos, X. (2018). Resilient first-order consensus and weakly stable, higher order synchronization of continuous-time networked multiagent systems. *IEEE Transactions on Control of Network Systems*. https://doi.org/10.1109/TCNS.2017.2696364

Lee, J., Bagheri, B., & Jin, C. (2016). Introduction to cyber manufacturing. *Manufacturing Letters*. https://doi.org/10.1016/j.mfglet.2016.05.002

Lee, J., Bagheri, B., & Kao, H.-A. (2015). A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters.*. https://doi.org/10.1016/j.mfglet.2014.12.001

Leveson, N. (2004). A new accident model for engineering safer systems. *Safety Science, 42*(4), 237–270. https://doi.org/10.1016/S0925-7535(03)00047-X

Leveson, N. (2011). *Engineering a safer and more secure world*. MIT.

Li, X., Wang, B., Liu, C., Freiheit, T., & Epureanu, B. I. (2020). Intelligent Manufacturing Systems in COVID-19 Pandemic and Beyond: Framework and Impact Assessment. Chinese Journal of. *Mechanical Engineering, 33*(1). https://doi.org/10.1186/s10033-020-00476-w

Li, Z., & Kang, R. (2015). Strategy for reliability testing and evaluation of cyber physical systems. In *2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). 2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*. https://doi.org/10.1109/ieem.2015.7385799

Li, Z., Shahidehpour, M., & Aminifar, F. (2017). Cybersecurity in Distributed Power Systems. In *Proceedings of the IEEE*. https://doi.org/10.1109/JPROC.2017.2687865

Lin, J., Yu, W., Zhang, N., Yang, X., Zhang, H., & Zhao, W. (2017). A Survey on Internet of Things: Architecture, Enabling Technologies, Security and Privacy, and Applications. *IEEE Internet of Things Journal, 4*(5), 1125–1142. https://doi.org/10.1109/jiot.2017.2683200

Linkov, I., & Kott, A. (2018). *Fundamental Concepts of Cyber Resilience: Introduction and Overview. In Cyber Resilience of Systems and Networks* (pp. 1–25). Springer International Publishing.

Lins, T., & Oliveira, R. A. R. (2020). Cyber-physical production systems retrofitting in context of industry 4.0. *Computers and Industrial Engineering, 139*, Article 106193. https://doi.org/10.1016/j.cie.2019.106193

Liu, J., Lu, X., & Wang, J. (2019). Resilience Analysis of DC Microgrids Under Denial of Service Threats. *IEEE Transactions on Power Systems*. https://doi.org/10.1109/TPWRS.2019.2897499

Lykou, G., Moustakas, D., Gritzalis, D., 2020. Defending airports from uas: A survey on cyber-attacks and counter-drone sensing technologies. Sensors (Switzerland). MDPI AG1–35. Doi: 10.3390/s20123537.

Ma, R., Shi, P., Wang, Z., & Wu, L. (2019). Resilient filtering for cyber-physical systems under denial-of-service attacks. *International Journal of Robust and Nonlinear Control, 30*(5), 1754–1769. https://doi.org/10.1002/rnc.4845

Mailloux, L. O., & Grimaila, M. (2018). Advancing cybersecurity: The growing need for a cyber-resiliency workforce. *IT Professional*. https://doi.org/10.1109/MITP.2018.032501745

Mailloux, L. O., Span, M., Mills, R. F., & Young, W. (2019). April). A Top Down Approach for Eliciting Systems Security Requirements for a Notional Autonomous Space System. In *2019 IEEE International Systems Conference (SysCon)*. https://doi.org/10.1109/syscon.2019.8836929

Marino, D. L., Wickramasinghe, C. S., Amarasinghe, K., Challa, H., Richardson, P., Jillepalli, A. A., et al. (2019). In *Cyber and Physical Anomaly Detection in Smart-Grids. 2019 Resilience Week (RWS). 2019 Resilience Week (RWS)*. https://doi.org/10.1109/rws47064.2019.8972003

Marquis, V., Ho, R., Rainey, W., Kimpel, M., Ghiorzi, J., Cricchi, W., et al. (2018). Toward attack-resilient state estimation and control of autonomous cyber-physical systems. In *2018 Systems and Information Engineering Design Symposium (SIEDS). 2018 Systems and Information Engineering Design Symposium (SIEDS)*. https://doi.org/10.1109/sieds.2018.8374762

Maurer, F., & Schumacher, J. (2018). Organizational Robustness and Resilience as Catalyst to Boost Innovation in Smart Service Factories of the Future. In *2018 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC)*. https://doi.org/10.1109/ice.2018.8436364

McDermott, T. A. (2019). A Rigorous System Engineering Process for Resilient Cyber-Physical Systems Design. In *2019 International Symposium on Systems Engineering (ISSE). 2019 International Symposium on Systems Engineering (ISSE)*. https://doi.org/10.1109/isse46696.2019.8984569

Miao, F., Zhu, Q., Pajic, M., & Pappas, G. J. (2018). A hybrid stochastic game for secure control of cyber-physical systems. *Automatica, 93*, 55–63. https://doi.org/10.1016/j.automatica.2018.03.01

Mohebbi, S., Zhang, Q., Christian Wells, E., Zhao, T., Nguyen, H., Li, M., et al. (2020). Cyber-physical-social interdependencies and organizational resilience: A review of water, transportation, and cyber infrastructure systems and processes. *Sustainable Cities and Society, 62*, Article 102327. https://doi.org/10.1016/j.scs.2020.102327

Moher, D. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *Annals of Internal Medicine, 151*(4), 264. https://doi.org/10.7326/0003-4819-151-4-200908180-00135

pp. Mokalled, H., Pragliola, C., Debertol, D., Meda, E., & Zunino, R. (2019). A Comprehensive Framework for the Security Risk Management of Cyber-Physical Systems. In Resilience of Cyber-Physical Systems (pp. 49–68). Springer International Publishing. Doi: 10.1007/978-3-319-95597-1_3.

Mörth, O., Emmanouilidis, C., Hafner, N., & Schadler, M. (2020). Cyber-physical systems for performance monitoring in production intralogistics. *Computers & Industrial Engineering, 142*, Article 106333. https://doi.org/10.1016/j.cie.2020.106333

Mouelhi, S., Laarouchi, M.-E., Cancila, D., & Chaouchi, H. (2019). Predictive Formal Analysis of Resilience in Cyber-Physical Systems. *IEEE Access, 7*, 33741–33758. https://doi.org/10.1109/access.2019.2903153

Murino, G., Armando, A., & Tacchella, A. (2019). Resilience of Cyber-Physical Systems: An Experimental Appraisal of Quantitative Measures. In *2019 11th International Conference on Cyber Conflict (CyCon). 2019 11th International Conference on Cyber Conflict (CyCon)*. https://doi.org/10.23919/cycon.2019.8757010

Naufal, J. K., et al. (2018). A2CPS: A Vehicle-Centric Safety Conceptual Framework for Autonomous Transport Systems. *IEEE Transactions on Intelligent Transportation Systems*. https://doi.org/10.1109/TITS.2017.2745678

Neema, H., Potteiger, B., Koutsoukos, X., Karsai, G., Volgyesi, P., & Sztipanovits, J. (2018). Integrated simulation testbed for security and resilience of CPS. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing. SAC 2018: Symposium on Applied Computing*. https://doi.org/10.1145/3167132.3167173

Neema, H., Potteiger, B., Koutsoukos, X., Tang, C., & Stouffer, K. (2018). Metrics-Driven Evaluation of Cybersecurity for Critical Railway Infrastructure. In *2018 Resilience Week (RWS). 2018 Resilience Week (RWS)*. https://doi.org/10.1109/rweek.2018.8473542

Nguyen, T., Wang, S., Alhazmi, M., Nazemi, M., Estebsari, A., & Dehghanian, P. (2020). Electric Power Grid Resilience to Cyber Adversaries: State of the Art. *IEEE Access, 8*, 87592–87608. https://doi.org/10.1109/access.2020.2993233

Pajic, M., Lee, I., & Pappas, G. J. (2017). Attack-resilient state estimation for noisy dynamical systems. *IEEE Transactions on Control of Network Systems*. https://doi.org/10.1109/TCNS.2016.2607420

Pajic, M., Weimer, J., Bezzo, N., Sokolsky, O., Pappas, G. J., & Lee, I. (2017). Design and implementation of attack-resilient cyberphysical systems: With a focus on attack-resilient state estimators. *IEEE Control Systems, 37*(2), 66–81. https://doi.org/10.1109/MCS.2016.2643239

Pajic, M., Weimer, J., Bezzo, N., Tabuada, P., Sokolsky, O., Lee, I., et al. (2014). Robustness of attack-resilient state estimators. In *2014 ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS). 2014 ACM/IEEE International Conference on Cyber-Physical Systems (ICCPS)*. https://doi.org/10.1109/iccps.2014.6843720

Panetto, H., Iung, B., Ivanov, D., Weichhart, G., & Wang, X. (2019). Challenges for the cyber-physical manufacturing enterprises of the future. *Annual Reviews in Control, 47*, 200–213. https://doi.org/10.1016/j.arcontrol.2019.02.002

Papagiannidis, S., Harris, J., & Morton, D. (2020). WHO led the digital transformation of your company? A reflection of IT related challenges during the pandemic. *International Journal of Information Management, 55*. https://doi.org/10.1016/j.ijinfomgt.2020.102166

Paridari, K., O'Mahony, N., El-Din Mady, A., Chabukswar, R., Boubekeur, M., & Sandberg, H. (2018). A Framework for Attack-Resilient Industrial Control Systems: Attack Detection and Controller Reconfiguration. *Proceedings of the IEEE, 106*(1), 113–128. https://doi.org/10.1109/jproc.2017.2725482

Park, S., Weimer, J., & Lee, I. (2017). April 18). Resilient linear classification. In *Proceedingsof th e 8th International Conference on Cyber-Physical Systems. ICCPS '17: ACM/IEEE 8th International Conference on Cyber-Physical Systems*. https://doi.org/10.1145/3055004.3055006

Pasqualetti, F., Dörfler, F., & Bullo, F. (2015). Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems. *IEEE Control Systems*. https://doi.org/10.1109/MCS.2014.2364725

Patriarca, R., Bergström, J., Di Gravio, G., & Costantino, F. (2018). Resilience engineering: Current status of the research and future challenges. *Safety Science, 102*, 79–100. https://doi.org/10.1016/j.ssci.2017.10.005

Patriarca, R., Di Gravio, G., Woltjer, R., Costantino, F., Praetorius, G., Ferreira, P., et al. (2020). Framing the FRAM: A literature review on the functional resonance analysis method. *Safety Science, 129*, Article 104827. https://doi.org/10.1016/j.ssci.2020.104827

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*, Article 105142. https://doi.org/10.1016/j.ssci.2020.105142

Pinzone, M., Albè, F., Orlandelli, D., Barletta, I., Berlin, C., Johansson, B., et al. (2020). A framework for operative and social sustainability functionalities in human-centric cyber-physical production systems. *Computers and Industrial Engineering, 139*. https://doi.org/10.1016/j.cie.2018.03.028

Ramadan, R. A., Aboshosha, B. W., Alshudukhi, J. S., Alzahrani, A. J., El-Sayed, A., & Dessouky, M. M. (2021). Cybersecurity and countermeasures at the time of pandemic. *Journal of Advanced Transportation, 2021*. https://doi.org/10.1155/2021/6627264

Ratasich, D., Khalid, F., Geissler, F., Grosu, R., Shafique, M., & Bartocci, E. (2019). A roadmap toward the resilient internet of things for cyber-physical systems. *IEEE Access, 7*, 13260–13283. https://doi.org/10.1109/ACCESS.2019.2891969

RELX (2019) RELX. Annual report and financial statements 2019 [Internet]. London: RELX. Available at: https://www.relx.com/~/media/Files/R/RELXGroup/documents/reports/annual-reports/2019-annualreport.pdf.

Russell, L., Kwamena, F., & Goubran, R. (2019, August). Towards Reliable IoT: Fog-Based AI Sensor Validation. 2019 IEEE Cloud Summit. 2019 IEEE Cloud Summit. Doi: 10.1109/cloudsummit47114.2019.00013.

Schneider, J., Dobie, S., & Ghettas, S. (2019). November). Translation of Process Safety to Cyber Incidents within the Emergency Management Arc. In *2019 IEEE International Symposium on Technologies for Homeland Security (HST). 2019 IEEE International Symposium on Technologies for Homeland Security (HST)*. https://doi.org/10.1109/hst47167.2019.9032932

Schoitsch, E. (2018). Smart systems everywhere - intelligence, autonomy, technology and society. Paper presented at the IDIMT 2018: Strategic Modeling in Management, Economy and Society - 26th Interdisciplinary Information Management Talks, 153-165. Retrieved from www.scopus.com.

Schoitsch, E. (2020). Towards a resilient society - technology 5.0, risks and ethics. Paper presented at the IDIMT 2020: Digitalized Economy, Society and Information Management - 28th Interdisciplinary Information Management Talks, 403-412. Retrieved from www.scopus.com.

Senejohnny, D., Tesi, P., & De Persis, C. (2018). A jamming-resilient algorithm for self-triggered network coordination. *IEEE Transactions on Control of Network Systems https: //*. https://doi.org/10.1109/TCNS.2017.2668901

Severson, T., Rodriguez-Seda, E., Kiriakidis, K., Croteau, B., Krishnankutty, D., Robucci, R., et al. (2018). June). Trust-Based Framework for Resilience to Sensor-Targeted Attacks in Cyber-Physical Systems. In *2018 Annual American Control Conference (ACC). 2018 Annual American Control Conference (ACC)*. https://doi.org/10.23919/acc.2018.8431909

Shen, J., & Feng, D. (2018). A game-theoretic method for cross-layer stochastic resilient control design in CPS. *International Journal of Systems Science*. https://doi.org/10.1080/00207721.2017.1406555

Shin, S., Lee, S., Burian, S. J., Judi, D. R., & McPherson, T. (2020). Evaluating resilience of water distribution networks to operational failures from cyber-physical attacks. *Journal of Environmental Engineering (United States), 146*(3). https://doi.org/10.1061/(ASCE)EE.1943-7870.0001665

Shintani, H., Aoyama, T., & Koshijima, I. (2017). Study on high resilient structures for IoT systems to detect accidents. *Journal of Disaster Research*. https://doi.org/10.20965/jdr.2017.p1073

Span, M., Mailloux, L. O., Mills, R. F., & Young, W. (2018a). Conceptual systems security requirements analysis: Aerial refueling case study. IEEE Access, 6, 46668-46682. Doi: 10.109/access.2018.28657366.

Span, M. T., Mailloux, L. O., R. Grimaila, M., & Young, W. B. (2018b). A Systems Security Approach for Requirements Analysis of Complex Cyber-Physical Systems. 2018 International Conference on Cyber Security and Protection of Digital Services (Cyber Security). 2018 International Conference on Cyber Security and Protection of Digital Services (CyberSecurity). Doi: 10.1109/cybersecpods.2018.8560682.

Sridhar, S., Hahn, A., & Govindarasu, M. (2012, January). Cyber attack-resilient control for smart grid. 2012 IEEE PES Innovative Smart Grid Technologies (ISGT). 2012 IEEE PES Innovative Smart Grid Technologies (ISGT). Doi: 10.1109/isgt.2012.6175567.

Stubs, M. (2018). October). Towards Emergent Security in Low-Latency Smart Grids with Distributed Control. In *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm). 2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*. https://doi.org/10.1109/smartgridcomm.2018.8587557

Sullivan, D., Colbert, E., & Cowley, J. (2018). August). Mission Resilience for Future Army Tactical Networks. In *2018 Resilience Week (RWS). 2018 Resilience Week (RWS)*. https://doi.org/10.1109/rweek.2018.8473522

Sun, Y.-C., & Yang, G.-H. (2018). Event-triggered resilient control for cyber-physical systems under asynchronous DoS attacks. *Information Sciences*. https://doi.org/10.1016/j.ins.2018.07.030

Sun, Q., Zhang, K., & Shi, Y. (2020). Resilient Model Predictive Control of Cyber-Physical Systems under DoS Attacks. *IEEE Transactions on Industrial Informatics*. https://doi.org/10.1109/TII.2019.2963294

Tao, F., Qi, Q., Wang, L., & Nee, A. Y. C. (2019). Digital twins and Cyber-Physical systems toward smart manufacturing and industry 4.0: Correlation and comparison. *Engineering, 5*(4), 653–661. https://doi.org/10.1016/j.eng.2019.01.014

Tomic, I., Breza, M., & McCann, J. A. (2019). Jamming-resilient control and communication framework for cyber physical systems. Living in the Internet of Things (IoT 2019). Living in the Internet of Things (IoT 2019). Doi: 10.1049/cp.2019.0132.

Umar, A., 2020. A software factory in the cloud for pandemics and other disasters – Initial results and future directions. In: Proceedings - 2020 IEEE Cloud Summit, Cloud Summit 2020. pp. 79–84. Doi: 10.1109/IEEECloudSummit48914.2020.00018.

Vachtsevanos, G., Lee, B., Oh, S., & Balchanos, M. (2018). Resilient design and operation of cyber physical systems with emphasis on unmanned autonomous systems. *Journal of Intelligent and Robotic Systems: Theory and Applications, 91*(1), 59–83. https://doi.org/10.1007/s10846-018-0881-x

Van Wyk, F., Wang, Y., Khojandi, A., & Masoud, N. (2020). Real-time sensor anomaly detection and identification in automated vehicles. *IEEE Transactions on Intelligent Transportation Systems, 21*(3), 1264–1276. https://doi.org/10.1007/s10916-020-1527-7

Venkataramanan, V., Hahn, A., Srivastava, A., 2019. Cyphyr: A cyber-physical analysis tool for measuring and enabling resiliency in microgrids. IET Cyber-Physical Systems: Theory and Applications Doi: 10.1049/iet-cps.2018.5069.

Venkataramanan, V., Hahn, A., & Srivastava, A. (2020). CP-SAM: Cyber-Physical Security Assessment Metric for Monitoring Microgrid Resiliency. *IEEE Transactions on Smart Grid*. https://doi.org/10.1109/TSG.2019.2930241

Venkataramanan, V., Srivastava, A., Hahn, A., & Zonouz, S. (2018). September). Enhancing Microgrid Resiliency Against Cyber Vulnerabilities. In *2018 IEEE Industry Applications Society Annual Meeting (IAS). 2018 IEEE Industry Applications Society Annual Meeting (IAS2018)*. https://doi.org/10.1109/ias.2018.8544667

Venkataramanan, V., Srivastava, A. K., Hahn, A., & Zonouz, S. (2019). Measuring and Enhancing Microgrid Resiliency Against Cyber Threats. *IEEE Transactions on Industry Applications, 55*(6), 6303–6312. https://doi.org/10.1109/tia.2019.2928495

Wadhawan, Y., Almajali, A., & Neuman, C. (2018). A comprehensive analysis of smart grid systems against cyber-physical attacks. *Electronics (Switzerland)*. https://doi.org/10.3390/electronics7100249

Wadhawan, Y., & Neuman, C. (2016). Evaluating Resilience of Gas Pipeline Systems Under Cyber-Physical Attacks. Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy - CPS-SPC '16. the 2nd ACM Workshop. Doi: 10.1145/2994487.2994488.

Wang, Y. (2018). Resilience Quantification for Probabilistic Design of Cyber-Physical System Networks. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems Part B: Mechanical Engineering* Doi, *10*(1115/1), 4039148.

pp. Wang, Y., Wang, Y., Liu, J., Huang, Z., & Xie, P. (2017). A survey of game theoretic methods for cyber security. Paper presented at the Proceedings - 2016 IEEE 1st International Conference on Data Science in Cyberspace, DSC 2016, 631-636. doi: 10.1109/DSC.2016.90 Retrieved from www.scopus.com.

Wedaj, S., Paul, K., & Ribeiro, V. J. (2019). DADS: Decentralized attestation for device swarms. *ACM Transactions on Privacy and Security*. https://doi.org/10.1145/3325822

Weerakkody, S., Ozel, O., Mo, Y., & Sinopoli, B. (2019). Resilient control in cyber-physical systems. *Foundations and Trends in Systems and Control, 7*(1–2), 1–255. https://doi.org/10.1561/2600000018

Weil, T., & Murugesan, S. (2020). IT Risk and Resilience-Cybersecurity Response to COVID-19. *IT Professional, 22*(3), 4–10. https://doi.org/10.1109/MITP.2020.2988330

Whelihan, D., Vai, M., Evanich, N., Kwak, K. J., Li, J., Britton, M., et al. (2017). October). Designing agility and resilience into embedded systems. In *MILCOM 2017–2017 IEEE Military Communications Conference (MILCOM). 2017 IEEE Military Communications Conference (MILCOM)*. https://doi.org/10.1109/milcom.2017.8170806

Woo, H., Yi, J., Browne, J. C., Mok, A. K., Atkins, E., & Xie, F. (2008). In *June). Design and Development Methodology for Resilient Cyber-Physical Systems*. https://doi.org/10.1109/icdcs.workshops.2008.62

Wu, X., Goepp, V., & Siadat, A. (2019). The integrative link between cyber physical production systems and enterprise information systems. Paper presented at the Proceedings of International Conference on Computers and Industrial Engineering, CIE, , 2019-October Retrieved from www.scopus.com.

Wu, C., Wu, L., Liu, J., & Jiang, Z.-P. (2020). Active Defense-Based Resilient Sliding Mode Control Under Denial-of-Service Attacks. *IEEE Transactions on Information Forensics and Security, 15*, 237–249. https://doi.org/10.1109/tifs.2019.2917373

Wuest, T., Kusiak, A., Dai, T., & Tayur, S. R. (2020). Impact of COVID-19 on Manufacturing and Supply Networks — The Case for AI-Inspired Digital Transformation. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3593540

Xu, Z., He, D., Vijayakumar, P., Choo, K. R., & Li, L. (2020). Efficient NTRU lattice-based certificateless signature scheme for medical cyber-physical systems. *Journal of Medical Systems, 44*(5). https://doi.org/10.1007/s10916-020-1527-7

Yang, B., Guo, L., Li, F., Ye, J., & Song, W. (2020). Vulnerability assessments of electric drive systems due to sensor data integrity attacks. *IEEE Transactions on Industrial Informatics, 16*(5), 3301–3310. https://doi.org/10.1109/TII.2019.2948056

Yang, K., Wang, R., Jiang, Y., Luo, C., Guan, Y., Li, X., et al. (2018). September). Enhanced Resilient Sensor Attack Detection Using Fusion Interval and Measurement History. In *2018 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS). 2018 International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS)*. https://doi.org/10.1109/codesisss.2018.8525941

Yoginath, S., Tansakul, V., Chinthavali, S., Taylor, C., Hambrick, J., Irminger, P., et al. (2019). November). On the Effectiveness of Recurrent Neural Networks for Live Modeling of Cyber-Physical Systems. In *2019 IEEE International Conference on Industrial Internet (ICII). 2019 IEEE International Conference on Industrial Internet (ICII)*. https://doi.org/10.1109/icii.2019.00062

Yong, S. Z., Zhu, M., & Frazzoli, E. (2015). December). Resilient state estimation against switching attacks on stochastic cyber-physical systems. In *2015 54th IEEE Conference on Decision and Control (CDC). 2015 54th IEEE Conference on Decision and Control (CDC)*. https://doi.org/10.1109/cdc.2015.7403027

Yuan, Y., Sun, F., & Liu, H. (2016). Resilient control of cyber-physical systems against intelligent attacker: A hierarchal stackelberg game approach. *International Journal of Systems Science*. https://doi.org/10.1080/00207721.2014.973467

Yuan, H., & Xia, Y. (2018). Resilient strategy design for cyber-physical system under DoS attack over a multi-channel framework. *Information Sciences*. https://doi.org/10.1016/j.ins.2018.04.082

Zhang, L. (2021). Emergency supplies reserve allocation within government-private cooperation: A study from capacity and response perspectives. *Computers and Industrial Engineering, 154*. https://doi.org/10.1016/j.cie.2021.107171

Zhang, P., Yuan, Y., Wang, Z., & Sun, C. (2019). July). A Hierarchical Game Approach to the Coupled Resilient Control of CPS against Denial-of-Service Attack. In *2019 IEEE 15th International Conference on Control and Automation (ICCA). 2019 IEEE 15th International Conference on Control and Automation (ICCA)*. https://doi.org/10.1109/icca.2019.8899933

Zheng, Z., & Reddy, A. L. N. (2017). April 2). Towards Improving Data Validity of Cyber-Physical Systems through Path Redundancy. In *Proceedings of the 3rd ACM Workshop on Cyber-Physical System Security. ASIA CCS '17: ACM Asia Conference on Computer and Communications Security*. https://doi.org/10.1145/3055186.3055189

Zhou, X., Li, Y., Barreto, C. A., Li, J., Volgyesi, P., Neema, H., et al. (2019). November). Evaluating Resilience of Grid Load Predictions under Stealthy Adversarial Attacks. In *2019 Resilience Week (RWS). 2019 Resilience Week (RWS)*. https://doi.org/10.1109/rws47064.2019.8971816

Zhu, Q., & Başar, T. (2012). A dynamic game-theoretic approach to resilient control system design for cascading failures. In *Proceedings of the 1st International Conference on High Confidence Networked Systems - HiCoNS '12. the 1st international conference*. https://doi.org/10.1145/2185505.2185512

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

Zio, E. (2018). The future of risk assessment. *Reliability Engineering and System Safety*. https://doi.org/10.1016/j.ress.2018.04.020

3.3. The research questions of the second part of the thesis

Following the literature review presented in the previous section and the identification of research gaps, a further step of narrowing down the research was taken to identify the second group of research questions. The purpose of this paragraph is to link each research questions to the findings of the extensive literature review.

- **A human centric cyber resilience**
  - RQ2.1: Which are the human factors involved in cybersecurity?
  - RQ2.2: How does each factor contribute as a weakness or opportunity to cyber resilience?

These research questions wants to investigate human's role in cybersecurity socio-technical systems. Despite much work being done in the field of cyber security, most of the attention seems to be focused on system usage and technical aspects. However, the increasing complexity and interconnectedness of today's systems does not allow for a simple inclusion of the human component in the system. Such a paradigm shift requires a new design of the human centric cyber security domain. To do this, it was deemed appropriate to start by identifying the characteristics that everyone has and does interact with the system. Specifically, the research will map human factors involved in cybersecurity and in which way those are linked to NIST cybersecurity framework functions. The objective is to investigate in which part humans are a threat or an opportunity for cyber resilience. The study will underline to scholars and practitioners how humans should be involved to be considered as a defensive or vulnerable agent in today's complex organizational scenarios.

- **Discussing the effectiveness of organizational cybersecurity outsourcing practices**
  - RQ3: Does the effectiveness of selected cybersecurity practices differ in the case of internally managed or outsourced cybersecurity processes?

The review has highlighted that the number of attacks has risen, and threats and incidents have become more sophisticated. There is a need to keep pursuing the shift from cybersecurity to cyber-resilience understanding the practical details of these events in real operational settings, and managing variability rather than simply trying to reduce it. In line with the need for a dynamic and fast-paced approach, the second objective of this thesis is to investigate whether outsourcing strategies can be a lever to increase organizational cybersecurity.

Literature states that organizations around the world are willing to focus on their core activities and grow their business which has led them to increasingly rely on external staff to manage specific aspects. The current situation requires qualified personnel, often difficult to have in-house to manage security operations. Using external experts can help find more qualified personnel, capable of understanding the dynamics of the domain considerably better. However, they are also questioning whether this strategy is increasing or reducing their threats and risks. Several organizations are frequently skeptical of the capabilities and extent of the cybersecurity service provider's solution, stating that when their technology and knowledge base are produced in-house, they assure superior system security. The danger of engaging with third parties and sharing corporate data is a typical justification used by organizations to justify not outsourcing security operations. The research aims to investigate the topic, trying to assess the situation in Italy, going to interview Italian SMEs cybersecurity experts and asking them about the decisions they have made in terms of cybersecurity managerial practices and their effectiveness. We will focus on those cybersecurity practices that are most found in organizational aspects such as: disciplinary processes, sanctions, norms, lessons learned.

- **Leveraging human-machine interaction for cyber resilience**
  - RQ4.1: How are conceptually grounded design elements for digital intelligent agents (DIAs)?
  - RQ4.2: How to enhance cybersecurity through DIAs?

Studies collected in the review addressed humans as flexible and able to rapidly judge and attack, stressing the importance of a continuous synergy among humans and machines to pursue cybersecurity effectiveness. Second, the review has also highlighted the importance of building resilience in cybersecurity. Specifically, there is an increase of interest in safety performance and this is underlying the fact that is no longer possible to separate cybersecurity from safety in critical infrastructure contexts. It is now important to revise processes in a perspective that combines security and safety. Moreover, it is suggested to consider as part of the solution the ability of humans involved in the socio-technical system to enhance and improve its cybersecurity and consequently its safety. We will refer to the term "security" as the protection of individuals, organizations, or assets from external threats. It refers to the practices and tools designed to protect cyber tools from external attacks. The term safety on the other hand, concerns a condition that allows one to be protected from which can cause harm and sometimes result in loss of life. As narrated in the background we are now faced with attack patterns that once the breach is successful can impact the safety of individuals.

In addition, during the years of this PhD, another paradigm shift involved industry, emphasizing the idea of Industry 5.0. Industry 5.0 decrease the emphasis on technology and assumes that the potential for progress is based on collaboration among humans and machines. Specifically, Industry 5.0 recognizes that human creativity and critical thinking cannot be replicated by machines. As such, ongoing innovation strives to optimize processes by delegating repetitive or predictable tasks to automation while also integrating human operators into production processes.

In this scenario, during the thesis work, a conceptual architecture and taxonomy for conversational agents to support production operators was developed. This solution was then used to deploy a conversational agent prototype to support cybersecurity. The objective is to assist firms in preparing to participate in Industry 5.0 and achieve better business outcomes while also building resilience to protect against evolving cyber threats.

4. A human centric cyber resilience

The "human factor" is considered the weakest link in creating secure digital environments, but human intuition can also be the solution to thwarting many cyber threats. Every software and security monitoring system requires human interpretation, and the latter can be trained and improved to make humans the primary tool of defense. In this chapter, to answer RQ2.1 and RQ2.2, after a thorough review of the literature, a classification of the human factors mainly involved in organizational cybersecurity was made to explain which is their impact, both positive and negative, on the security of an information system. Once identified, they were then incorporated into the authoritative NIST framework, so that it could be enriched with the human dimension that provides important additional information to be able to ensure the integrity of the system. Human factors have been linke d to the 5 functions of NIST: IDENTIFY, PROTECT, DETECT, RESPOND and RECOVER. In particular, the research draws conclusions highlighting which human factors are most involved in the cyber security functions and how the organization must proceed to integrate the human factor into its cyber practices.

4.1. Human Factors in cybersecurity

The field of human factors seeks to improve the interaction between people and technology. The International Ergonomics Association defines human factors as the *"scientific discipline concerned with the understanding of the interaction among humans and element of the system"*[88]. Human factors have been analyzed in various fields, especially heal and aviation industries which have extensive work in this discipline [89]. Among the most famous classifications of human factors is for instance "The Dirty Dozen" proposed by Dupont in 2009 for the aviation sector [90] and taken up by other fields such as healthcare and aviation [91], [92], and not least cybersecurity [93]. Focusing on human factors and cybersecurity research, various classifications, ontologies, or just thoughts on what aspects of human character most impact cybersecurity have been proposed over the years. Specifically, the characterization of human factors, which includes human behavior, is necessary to understand how the actions of users, defenders (IT personnel), and attackers influence cybersecurity risk. Among the most comprehensive works is that of [94]. The authors propose a human factors trust ontology named Human Factors Ontology (HUFO). The researchers, by focusing on the notion of trust, create and enumerate risk characteristics and relate them to human factors. These are then broken into three main categories of the attacker, defender, and user which interact with computer networks. Their objective is to propose an application of HUFO as a support tool for risk assessment and risk prioritization in cyber operations.

In another extensive work [93], the authors conducted a systematic review of the most important researches on human factors and phishing. Specifically, they use the well-known Dupont factors as classification categories and detail the misbehaviors of individuals related to phishing phenomena.

Another contribution [95] examines how risk-taking preferences, decision-making styles, demographics, and personality traits influence the security behavior intentions of device securement, password generation, proactive awareness, and updating.

Finally, an interesting work is carried out by the Chartered Institute of Ergonomics and Human Factors [96], which has compiled a list of risky human behaviors which is then correlated to cybersecurity issues and vulnerabilities.

Other contributions have instead explored the correlation of human traits, specific cyber security behavior intentions, or impact on a person's adherence to cybersecurity procedures, rules, and practices. These studies have focused on quite specific aspects of personality or types of attacks such as Social E attacks. This is the case, for example, with the work of William et al. [97] exploring the susceptibility of workers to phishing attacks and how they are affected when they receive an email in which the authority of the sender or the urgency of the task is determined. Similarly, Ubelacker and Quiel [98] examined the link between susceptibility to social engineering attacks and the big five personality traits. Finally, Hadlington [99] analyzes the correlation between impulsivity and Internet addiction toward risky cybersecurity behaviors.

4.2. Human Factors in Cybersecurity Frameworks

In the previous paragraph, a selection of relevant studies aimed at enumerating the factors involved in cybersecurity was reported. In this section, the author will present the current state of research in terms of cybersecurity models and frameworks that propose guides and best practices for fruitful integration and management of human behaviors in the presence of cyber threats. These models are being considered for future research steps pursued.

Organizations that implement strong technological security practices often still do not pay enough attention to the human sources of vulnerability. Combining education and technology ensures that mistakes, even when made, do not lead to the demise of the organization. Undoubtedly, the human factor is a scientific area that is underutilized and undervalued in cybersecurity [94]. Human involvement in information security is too valuable for organizational leaders to continue to ignore the importance of analyzing human behavior in information security [100].

[101] argued that empirical and theoretical research on human aspects of cybersecurity based on the volume of human error-related incidents should be increased to find ways to improve cybersecurity. In their article, they indicated that further research on cybersecurity and the quantification of human factors is needed to

develop an effective security framework. They suggest that the next step would be the development of a framework comprehensive of a range of human aspect tasks that provide or are intended not to negatively affect cybersecurity posture. Cybersecurity frameworks only marginally talk about the human factor, as they usually mainly talk about the application of corporate policies.

Among the most comprehensive and useful for conducting a corporate security assessment is that devised [102]. The paper introduces a human-centered approach to threat modeling, titled STRIDE-HF, which extends the existing threat modeling framework STRIDE by linking the STRIDE elements to the Dirty Dozen of Dupont. The STRIDE method is a mnemonic for six types of security threats. It forms the basis for the theoretical model STRIDE-HF (Spoofing, Tampering, Repudiation, Information Disclosure, Denial of Service, and Elevation of Privilege - Human Factor). Its elements are defined as follows:

- *(S): Spoofing: using someone else's credentials to gain access to otherwise inaccessible assets.*
- *(T): Tampering: changing data to mount an attack.*
- *(R): Repudiation: occurs when a user denies performing an action, but the target of the action has no way to prove otherwise.*
- *(I): Information disclosure: the disclosure of information to a user who does not have permission to see it.*
- *(D): Denial of service: reducing the ability of valid users to access resources.*
- *(E): Elevation of privilege: occurs when an unprivileged user gains privileged status.*
- *(HF): "Dirty Dozen" Human Factors: Twelve of the most common human factor-related errors, which may lead to aviation-related accidents or incidents.*

Their research highlighted the need of considering human factors as a type of threat to system security in a socio-technical world. The STRIDE-HF framework provides a way for security analysts to consider human factors behavior and assess the types of security breaches that could result. For example, if a user shares a password, this could result in an elevation of privilege where a user unknowingly disables certain settings, creating a vulnerability. The main difference that STRIDE-HF offers compared to traditional threat modeling methods is that it takes a "reverse" approach to classify threats that may affect the security of a system from the user's perspective rather than that of an attacker. The authors suggest that future iterations of STRIDE-HF may include additional human factors as well as include more psychological-based principles. Finally, the research also stressed the importance of extending the work by going into greater detail on human factors-related aspects so that more specific remedies can be offered for safety and security critical situations.

Another interesting work is that of [103] which partially complements the first study presented. While the previous framework allowed security analysts to consider human factors' behavior and assess the types of security breaches that could occur, the goal of this framework is to develop measures that take into account the strengths and weaknesses of people and devices and test them in real systems to improve cybersecurity. They describe cybersecurity as a system state. A system might be somewhat secure, unsecure, or ambiguous. Any action taken by an actor can have an impact on the system's safety. They describe three types of players in the framework: IT Providers, target entities, and threat entities. Architects, software developers, and interface designers are examples of IT providers who work on the creation of digital systems and products. System administrators, legal counsel, managers, and directors are examples of target entities involved in the utilization of digital systems and goods. Threat entities are actors who compromise system security, such as criminal organizations.

Their framework provides a set of behaviors that can improve cybersecurity but does not identify the determinants that cause specific failure, which is the inverse of the previous framework and thus does not provide enough support to help organizations protect their systems from threats before they occur.

Finally, the work of the Chartered Institute of Ergonomics and Human Factors [104] which presents a practical human factors (HF) framework and checklist that can be used to enhance cyber security. The Human Affected Cyber Security (HACS) framework incorporates risky behaviors, causes, and solutions. Their solution is designed to support HF practitioners, human reliability analysts, and cyber security professionals who would like to investigate more about the contribution of the human element.

The framework provides a structure to capture human-related cybersecurity vulnerability in organizations. causes and mitigating solutions. It can be used proactively, as part of a cyber security risk assessment, or retrospectively, in an incident investigation. The framework shows the interactions among aspects such as organizational culture, ways of working, and individual characteristics which cause risky behaviors. The framework should be seen as a starting point for HF practitioners to adapt as technology and working practices evolve, and as new research is published.

Therefore from what is observed in the literature, the role of human factors in cybersecurity is becoming widely researched. Moreover, the increasing interest in cybersecurity human factors has led to the theorization of multiple reference cybersecurity human factors frameworks, taxonomies, classification, etc. However, most studies have focused on investigating human factors as contributions to human errors. Most frameworks either provide an overview of the human dimension that is at the root of errors in cybersecurity, or they provide a set of human behaviors that can increase the security of the system. As far as the authors know, there is still no framework that is inclusive of both views: on the one hand the human aspects that generate vulnerability and on the other hand the strength of behaviors capable of reacting and intercepting cyber threats. First, there is a lack of structured and synthesized knowledge that scholars and practitioners can use to develop further studies on information security culture, cyber training programs, and investment decisions. Second, there is a need for a shared, user-friendly classification that brings together all the human factors involved in cybersecurity and links them with a set of best practices that encompass the system people, process, and technology. In addition, all the reviewed studies report attention to a dimension of cybersecurity that as stated earlier is not exhaustive of what systems need today. This thesis includes the human factor, as a mitigator and vulnerability, within a program aimed at increasing cyber resilience.

Therefore, the next paragraphs will collect and list through the most up to date research on the human factors involvement in cybersecurity. Specifically, it is intended to report in which way each of these factors is a driver or a barrier to cybersecurity. Moreover, the list will be integrated with the well-known NIST cybersecurity framework by accounting humans as a risk factor and a risk mitigator for cyber resilience.

Three sub-research questions are thus identified and will be answered in the following sections:

- Which are the human factors involved in organizational cybersecurity?
- In which way does each human factor contribute as a weakness or opportunity to organizational cybersecurity?
- Which role do the defined human factors play in the functions and categories of the NIST framework?

4.3. Human factors in cybersecurity: weakness or opportunity?

4.3.1. Human factors collection

In this first phase, an attempt was made to answer the question that sought to identify the factors involved in cybersecurity. To do this, an analysis of the literature was conducted. Several search keys were structured by combining generic terms such as human factors and human error and then more specific terms related to Dupont's human factors. It was chosen to retain this categorization of factors both because of its use noted in the cybersecurity field [93] and because sufficiently broad to include multiple subfactors useful in describing all critical issues. Dupont defines the twelve factors as the "Dirty Dozen" [90]. The name was decided to indicate their negative connotation. In his work, the author states the importance of identifying human factors that can contribute to error, regardless of scope. Moreover, the author underlines the optimal approach for reducing human error: identify human factors;

introduce human factors training; provide a work environment resistant to human error. The list outlined by Dupont considers the following factors:

- *Lack of Communication: People not communicating with each other within a working and/or online environment*
- *Complacency: A feeling of self-confidence that can lead to lack of awareness of potential dangers*
- *Lack of Knowledge: Not having specific knowledge and enough experience, which can lead to poor decisions.*
- *Distraction: When users' attention has been taken away from the task that they are required to do.*
- *Lack of teamwork (Trust): Not providing enough support toward a group of people, co-workers, and so forth, who rely on your support*
- *Fatigue: A physiological reaction resulting from prolonged periods of work and stress.*
- *Lack of resources: Not having enough resources (e.g., time, tools, people) to complete a task.*
- *Pressure: Pressure to meet a deadline interferes with our ability to complete task correctly.*
- *Lack of assertiveness: Not being able or allowed to express concerns or ideas.*
- *Stress: Acute and chronic stress from working for long periods or the other demanding issues such as family or financial problems.*
- *Lack of awareness: Not being aware of what happens in the surrounding (working or online) environment, often leading to an unconscious disconnection from what others are doing.*
- *Norms (Policies): workplace practices that develop over time, which can then influence other behavior.*

However, considering the different contexts in which they operate these need to be readjusted. First, when discussing human error in cybersecurity, it refers to unintentional actions, or lack of action, by users that cause or allow a security break to take place. The authors will not consider security-related human errors those related to software engineering or IT. Second, Dupont factors again reflect only the negative connotation of human factors. This research will take these factors as a reference by considering only the neutral meaning of their factors (e.g. communication and not lack of communication). According to these criteria, the final search strings used for our search were the following:

- *"human factors" and ("cybersecurity" or "cyber security")*
- *"human error" and ("cybersecurity" or "cyber security")*
- *"human behavior"and ("cybersecurity" or "cyber security")*
- *"x" and ("cybersecurity" or "cyber security")*

where "*x*" is substituted by either "*communication*" or "*complacency*" or "*knowledge,*" or "*distraction,*" or "*teamwork*", or "*fatigue*", or "*resources*", or "*pressure*", or "*assertiveness*", or "*stress*", or "*awareness*", or "*norms*" for a total of 14 search strings.

The databases used for the search is Scopus. Once the core articles were identified, the papers cited and those that cited them were then also explored.

4.3.2. Human factors classification

In order to achieve an easy-to-use classification, it was decided to decline each factor on an organizational dimension. Each factor is declined on an organizational dimension: individual, team, or work. This division is made by categorizing human error as caused by internal (individual), and external (organizational and work) factors relying on the research on Performance Influencing Factors (PIFs) formerly also called Performance Shaping Factors (PSF) [105]. PIFs address what people are being asked to do (the Work and its characteristics), who is doing it (the People and their competence), and where they are working (the Organization and its attributes). For each factor, thanks to the analysis of the literature, subfactors that represent help or criticality in cybersecurity management are proposed. To date, despite the wide use of PIFs in human reliability studies [106] there is a lack of studies that applies PIFs to cyber security. However, similar classifications have emerged underlining how individual, work, and organizational aspects might influence the actions and behaviors of humans [107]–[109].

PIFs literature affirms that tasks should be designed in accordance with human factors principles to address people's limitations and strengths. Since discrepancies between job requirements and people's capabilities increase the potential for human error, matching the job to the person ensures that the person is not overworked and makes the most effective contribution to the organization. In general, the work environment should address both the physical and psychological dimensions of employees to help them reach their full potential. The physical dimension refers to the design of the workplace and the workload itself. The mental dimension includes the management of communications (both for positive events, such as an employee's success in thwarting a cyberattack, and for negative events, such as the loss of information due to a phishing attack), but also the awareness that managers can convey about the importance of their activities and the risks in the cyber world.

Some examples of work errors can be:

- Poor work planning, leading to high work pressure
- Inadequate responses to previous incidents
- Deficient coordination and responsibilities
- Clarity of signs, signals, instructions, and other information)
- Procedures inadequate or inappropriate

- Preparation for a task (e.g. permits, risk assessments, checking)
- Working environment (noise, heat, space, lighting, ventilation)
- Time available/required

When talking about people, it can be said that they primarily have different characteristics, such as skills, personality, habits, and work experience. These characteristics can be a strength or a weakness for the individual, depending on the work required. In fact, work performance can be positive or negative depending on the individual. While some human characteristics, such as skills and experience, can be changed and improved, others, such as personality, are nearly impossible to change. Some examples of human factors in a work environment are:

- Low skill and competence levels
- Tired staff
- Bored or disheartened staff
- Physical capability and condition
- Work overload/underload

Finally, the organization can significantly influence people's behavior. If we talk about organizational factors, corporate culture can be mentioned as before. Based on this, individuals know what behavior to display and maintain in the organization, and how to proceed. The errors that might arise owing to organizational factors are typically governed by concerns that are upstream of the organization and do not depend on the persons conducting the activity. The business plan of an enterprise should incorporate the establishment of a successful information security-focused organization. This entails developing rules that foster a culture in which workers are unwilling to breach information security protections in order to do their tasks [110]. Employees in an enlightened security culture actively raise their awareness and concern about the importance of information security and recognize that it is a component of everyone's role, not just those with information technology responsibilities. Among the various errors we can mention:

- Level and nature of supervision
- Peer pressure
- Clarity of roles and responsibilities
- Effectiveness of organizational learning (learning from experiences)
- Poor work planning, leading to high work pressure
- Management based on one-way communications

Another useful and widely used classification in the context of cybersecurity and human factors is the Knowledge-Attitude-Behaviour (KAB) model defined and used by [111]–[113] . The approach investigates workers' "Knowledge" (K) of policy and procedures; “Attitudes” (A) towards cybersecurity policy and Procedures and self-reported “Behavior’s” (B). This model is often associated with factors such as individual factors, organizational factors, work factors, and technology factors [108] which recall what was theorized by PIFs. It is observed how such contextual factors influence the knowledge, attitude, and behavior of individuals thus generating positive or negative action on the system in which they operate. For example, from an individual perspective, a specific collection of individual elements, such as subjective norms; beliefs in the perceived repercussions of an action or real understanding of the cybersecurity issue, may shape attitudes. At the same time, employees' attitudes might allow additional social and organizational elements to impact them, such as social standards, ethical dilemmas, and different levels of behavioral control experienced by the employee. On the other hand, well-informed and trained employees reduce the occurrence of unintentional and non-deliberate actions that constitute a violation of cybersecurity rules, and they play an important role in reducing information security risks and protecting the organization's critical assets and valuable intellectual property [114].

From a business and organizational perspective, one can observe how the management and communication of policies, standards, and processes impact cybersecurity. On the one hand, organizations have these formal aspects to guide employees in keeping the system secure and expect employees to comply with them. On the other, these do not regulate the human aspect, which instead takes shortcuts in the name of improving efficiency or simply helpfulness, even if it involves implementing a breach [115], [116]. Such procedures succeed in being effective and thus positively influencing employees only if they are comprehensively communicated, and shown as resources for action and not as a way to regulate human behavior.

4.3.3. Human factors: drivers and barriers

Once the collection of factors was concluded, this stage proceeded to evaluate in which way each human factor contributes as a weakness or opportunity to organizational cybersecurity. Recent literature has begun to show how a given factor can be told either positively or negatively with respect to a cyber threat. For example, the "communication" factor is researched on the one side as “a lack of communication is an origin for errors” [103], on the other side “Communication between humans and technology and between different human actors is essential to ensure cyber security” [117]. For this reason, the present research aims not only to provide information on how human factors impact cybersecurity but also to characterize them by explaining in the presence of which "drivers" the human factor is effectively exploited to improve cybersecurity by making it an opportunity and which "barriers" turn it into a weakness, increasing the vulnerability of the system. Two categories have been created: "Drivers to cybersecurity" and "Barriers to cybersecurity". Moreover, to allow a homogeneous and comprehensive reading, these circumstances in which the human factors are a driver or a barrier are divided towards the PIF's introduced above to better identify their impacts. A detailed description of it is given below.

1. **Communication**:

A process by which information is exchanged between individuals through a common system of symbols, signs, or behavior. In the context of cybersecurity means people communicating with each other within a working and/or online environment. Communication between humans and technology and between different human actors is essential. It starts with communication within work teams, but also between different organizational departments and with humans outside the organization, such as customers or cooperating partners. Communication among industries and governments across national borders helps in identifying, and targeting cybersecurity threats and increasing cyber resilience. The communication about cybersecurity should not only include experiences of handling threats and adverse events, but also the sharing of success stories that can be beneficial in deriving lessons for improving resilience [118].

Drivers to cybersecurity:

Individual:

- Success stories. To increase awareness, understanding, and use of cybersecurity tools it is suggested to publish and communicate success stories and praising employees who spot attacks or alert colleagues. [119].
- Soft skills. Soft skills are all those skills that have to do with interpersonal and communication skills such as critical thinking, problem-solving, public speaking, professional writing, teamwork, digital literacy, leadership, professional attitude, work ethic, career management, and intercultural fluency. Communication is one of the most essential soft skills to train on to avoid internal communication barriers, especially when facing an attack [120].

O*rganizational:*

- Promotes the feeling of belonging to a group. Informal communication plays a key role in collaboration within companies. With informal communication, it is easier to create cohesive work groups. In these groups, knowledge exchange and collaborations give to the team members the feeling of belonging to a group [121]. In this way, communication within the organization will improve thanks to the sense of belonging and therefore the individual will perceive the work group as a place where they can work together to manage an attack.

- Clarification of role and responsibilities. Communication of plans, policies, and role expectations should be clarified. The purpose is to guide and coordinate work activities and make sure people know what to do and how to do it. Each employee needs to understand what duties, functions, and activities are required for the job and what results are expected. Even an employee who is highly competent and motivated may fail to achieve a high level of performance in presence of miscommunication [122].

Work:

- Internal communication. Internal communication is a part of the management process, through which information is shared, collected, and distributed, as to ensure employee understanding of the organization's goals and objectives [123]–[126]. Internal communication plays a key role in keeping the employees informed about the organization's plans, vision, and ideas. Moreover, it encourages them to participate in the decision-making processes, as well as promotes employee feedback and peer learning and this reflects in a more productive work environment. Specifically fruitful internal communication in the context of cybersecurity is related to communicating efficiently the new norms adopted by the organization.
- Strong Feedback system. Employee engagement levels can be regularly surveyed to receive feedback and identify motivating factors. Such survey results should prompt managers to create clear, measurable, accountable, and time-bound action plans [127]. By making employees accountable for their performance, they'll be stimulated to establish effective communication and better performance which should result in good attack response and trust.

Barriers to cybersecurity:

Individual:

- Fear. Fear is an individual barrier to communication in the context of cybersecurity. "Fear" to communicate a mistake to one of our superiors for the negative consequences, or to communicate with one of our colleagues because we have "fear" of their judgment. Specifically, the theme of fear connects with that of lessons learned in post-cyber attack recovery. To learn from events, and to be able to react quickly, reporting a cyber incident is essential. However, reporting is unlikely if employees or customers fear negative consequences such as blaming, shaming, financial loss, prosecution, or job loss [117].
- Cognitive-type fatigue. The limit of cognitive resources an individual can devote to security concerns, that could result, for instance in employees tuning out cybersecurity-related communications [128].

- Interpersonal conflict. Interpersonal conflict refers to the representation of incompatibility, disagreement, or difference between two or more interacting individuals [129]. Conflict can impair communication between parties in the workplace. It can lead people not to cooperate or prevent the parties from addressing real issues or problems [130].

Organizational:

- Management based on one-way communication. Managers should encourage two-way dialogue. Clear and consistent communication engages the workforce. In addition, sharing authority with staff through participatory decision-making enhances a sense of ownership. This connects to an improvement in risks related to phishing attacks by reducing distractions often related to a sense of anxiety toward the sending authority [127].
- Development of poor cyber security policy. The development of ineffective cybersecurity policies and state frameworks makes it impossible to compromise between safeguarding the security of the information system and providing access to these channels as easily and directly as possible [131].
- Poor work planning. Work planning can help a lot to develop a great strategy to protect the system of an organization. The planning approach ought, to begin with a cybersecurity risk assessment that identifies important business goals, crucial IT resources for achieving those goals (clarification of roles), implementation of effective communication channels and prospective cyberattacks as well as how probable the attacks are to happen and what kind of business impact they could have. If there is poor work planning the organization may experience communication problems causing vulnerabilities in cybersecurity [132].

Work:

- Lack of incentives. Employees who demonstrate greater job engagement should receive financial and non-financial advantages. According to several management theories, workers are more likely to put more effort into their work when they receive praise and recognition. Performance and employee incentives should be directly related [127]. A lack of incentives can lead to a lack of participation and feeling to protect and safeguard the organization. Risks such as leakage of information, passwords, and sensible data increase.

Table 3 - Human Factor: Communication

Human Factors	PIFs	Drivers to cybersecurity	Barriers to cybersecurity
Communication	*Individual*	Success Stories Soft skills	Fear Cognitive-type fatigue Interpersonal conflict
	Organizational	Promotes the feeling of belonging to a group Clarification of role and responsibilities	Management based on one-way communication Development of poor cyber security policy Poor work planning
	Work	Internal communication Strong feedback system	Lack of incentives

2. **Complacency:**

A feeling of being satisfied with yourself or with a situation, so that you do not think any change is necessary. Physical security complacency and IT security complacency both refer to maintaining the status quo in the face of changing requirements and threat scenarios. When it comes to achieving better results, complacency might indicate overconfidence or disinterest. Alternatively, it can indicate a certain degree of desensitization to online dangers. Some researchers have suggested that employee overconfidence and complacency can explain the negative relationship between cybersecurity training frequency and workplace behaviors [107], [128], [133].

Drivers to cybersecurity:

- As can be guessed from the definition, complacency is a negative human factor with which no cybersecurity benefit is associated according to the literature; for this reason, it cannot be considered a driver at any level (individual, organizational, and work).

Barriers to cybersecurity:

Individual:

- Lack of experience. People's apparent complacency is mainly due to their lack of direct experience related to a significant cyber incident that disrupted a critical service. [134].
- Lack of strong desire. The lack of a desire to maintain an adequate degree of success within the workplace is a prerequisite for complacency [135].
- Overtrust in cybersecurity devices. Employees frequently lack a broad comprehension of how cybersecurity appliances operate. This causes individuals to overestimate themselves and become distracted, believing that the devices can create an effective line of defense without their supervision [136]. When human operators observe automation complacently, they are less inclined to manually take control when their intervention is required. The development of misplaced trust may be linked to complacency [137].

Organizational:

- Lack of autonomy. Giving employees greater work autonomy will allow them to choose the best method to complete their tasks and increase the quality of results [127]. Empowering the employee with responsibilities will result in greater attention and supervision of the work performed, avoiding a superficial attitude. In addition, greater autonomy speeds up the acquisition of experience and counteracts complacency

Work:

- Shifts away from self-responsibility. Many computer users experience a feeling of insecurity. This also leads to a shift away from self-responsibility regarding security procedures. Users rely on others (appliances or people) convinced that they will be in charge of system security without the need for personal involvement [138].

Table 4 - Human Factor: Complacency

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Complacency	***Individual***		Lack of experience Lack of strong desire Overtrust in cybersecurity devices
	Organizational		Lack of autonomy
	Work		Shifts away from self-responsibility

3. **Knowledge:**

   Facts, information, and skills acquired through experience or education; the theoretical or practical understanding of a subject. Given the multidisciplinary nature of organizational cybersecurity, a knowledge-based perspective seems essential.

   Organizational cybersecurity management must converge on knowledge management. This acts by seeking to safeguard both intellectual property and ensure business continuity. For this reason, it is observed that a knowledge-based perspective on cybersecurity and its management would have a direct impact on the dynamics between individuals, technology use, and trust [139].

   Drivers to cybersecurity:

   Individual:

   - Motivation. Two of the most important elements in the effectiveness of cybersecurity education and training are user engagement and motivation. An employee who is committed and motivated is more inclined to increase his or her knowledge and thereby be better prepared in the face of cyber threats [140]–[142].
   - Employee satisfaction. According to [143], job satisfaction is characterized by how much people like or dislike their occupation and the nature of their tasks. Low absenteeism, employee turnover, and improved job performance are all results of high employee satisfaction. Higher skills and better performance will increase the level of employee satisfaction [144]. According to [137], job satisfaction is characterized by how much people like or dislike their occupation and the nature of their tasks. Low absenteeism, employee turnover, and improved job performance are all results of high employee satisfaction. Higher skills and better performance will increase the level of employee satisfaction [138]. In order to improve performance and, consequently, personal satisfaction, it is necessary to be aware of how cybersecurity appliances work and to have extensive knowledge to perform one's job with less pressure and risk of an accident.

   Organizational:

   - Training. Standards, such as NIST [60] and academic researchers [145], [146], highlight the necessity of training as a way to increase knowledge for more efficient cybersecurity. Consequently, many approaches exist to increase security knowledge and awareness [147], [148]. Specifically, a recent Microsoft report recognizes that humans are often considered the weakest link in security but *"with the training and education they can also be the first line of defense"* [149].

- Cyber Hygiene (CH). Cyber hygiene is a relatively recent concept that emphasizes the importance of social and human factors in reducing vulnerabilities and the danger of attacks and breaches. It encompasses a set of practices that organizations and individuals regularly perform to maintain the health and security of users, devices, networks, and data. Understanding and using good cyber hygiene practices requires user training and awareness [150].

Work:

- Information sharing. Sharing of success stories that can be beneficial in deriving lessons for improving resilience [117].
- Cyber drill. The cyber drill is a training process that simulates a cyber attack on employees whose work is related to cyber incident response [151].

Barriers to cybersecurity:

Individual:

- Lack of knowledge sharing. When people's motivation, fear, and trust are compromised they are prevented from sharing knowledge [152]. Organizations should reward, motivate, and encourage employees to guarantee that knowledge transfer takes place. Regarding technical aspects, limitations such as lack of knowledge about the usability of platforms, training on their use, excessive information, and poor understanding of social media emerge.
- Overtrust in cybersecurity devices. Employees frequently lack a basic knowledge of how cybersecurity equipment operates. This causes them to overtrust them, and while they are confident in their capacity to safeguard servers, they do not gain the knowledge required to comprehend their usage. [136].

Organizational:

- Absence of a security-oriented organization. Developing a security-oriented organization is fundamental. Employees will deliberately expand their understanding of the value of information security and their concern for it in this security-conscious culture, realizing that this is a component of everyone's employment, not just those with roles and responsibilities related to information technology [153].
- Lack of management skills. among management skills, related to cybersecurity it is worth mentioning:
    - Risk management
    - Identity and access management;
    - Asset, change, and configuration management;
    - System administration;

- Workforce management;
- Cyber-security program management;
- Supply chain and external dependencies management;
- Evaluation of policies effectiveness;
- Project planning; [154]

If one of these aspects is not managed properly, not negligible cyber issues can arise.

Work:

- Lack of time. According to the European Union Agency for Cybersecurity [155], manufacturers and other organizations using a wide range of Internet of Things applications often do not have time to train staff adequately causing poor cybersecurity knowledge and exposing organizations to potential risks.
- Absence of cognitive anchoring. Without a cognitive anchor, the context and relationships in which individuals operate generate beliefs that guide their behaviors. This background limits the ability to respond and relies on beliefs and presumed knowledge as the mode of response. [139].
- Lack of Operator learning. Employee learning focuses on the process of individual skill acquisition. Human capital represents the value of the skills of all employees as a result of the information and skills that employees in the organization have learned [156]. Such capital is a crucial asset [157]. In cybersecurity, increasing staff knowledge can prevent cyber attacks involving unskilled employees by generating a security network with those who are more prepared [144].

Table 5 - Human Factor: Knowledge

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Knowledge	*Individual*	Motivation Employee satisfaction	Lack of knowledge sharing Overtrust in cybersecurity devices
	Organizational	Training Cyber hygiene	Absence of a security-oriented organization Lack of management skills
	Work	Information sharing Cyber drill	Lack of time Absence of a cognitive anchoring Lack of operator learning

4. **Distraction:**

Distraction prevents individuals from concentrating. Such distractions have cognitive interference consequences beyond the scope of assigned tasks. Distractions impair our ability to process a message and thus we use heuristics to simplify information. Therefore, perturbations in our cognitive processes impact not only our ability to make decisions but also the ease with which we are persuaded by new information [158].

Drivers to cybersecurity:

As can be guessed from the definition, distraction is a negative human factor with which no cybersecurity benefit is associated; for that reason, it cannot be considered a driver at any level (individual, organizational, and work).

Barriers to cybersecurity:

Individual:

- Personal characteristics. Studies have shown that personal characteristics such as the ability to maintain concentration, personal fatigue, and unique personality traits like being conscientious or nervous could affect operators' perception of distractions at the workspace and their stress levels [159].
- Overtrust in cybersecurity devices. Employees, often do not have a broad understanding of how appliances devoted to cybersecurity work. This leads them to overtrust them and become distracted, confident that the devices can form an effective line of defense autonomously, without their supervision [136].

Organizational:

- Presence of Internet ads. It seems reasonable to predict that when ads are present during an online task, people very well may be more likely to rely on stereotypical knowledge to complete the task than if such ads are not present. The intrusion of Internet ads represents a cognitive distraction and consumes a portion of the resources available to devote to other cognitive tasks [158].

Work:

- High workload. When the worker experiences mental stress conditions or a high workload, he or she is more prone to human errors generated by distraction [116].
- Working from home. The risk of a remote worker accidentally introducing malware links into the company's computer network increases due to the increased distractions that are reported to have with home-based work and the reduced amount of technical protections [159].

- Workspace conditions. The temperature, noise level, size of the workspace, the adjustability of the furniture, the colors on the walls, and the cleanliness of the workspace could all potentially cause workspace distractions [159].

Table 6 - Human Factor: Distraction

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Distraction	***Individual***		Personal characteristics
			Overtrust in cybersecurity
	Organizational		Presence of Internet ads
	Work		High workload
			Working from home
			Workspace conditions

5. **Teamwork:**

Teamwork is the activity of working in groups with other people. Several studies have highlighted how teamwork can grow organizations' skills. Even in the cyber domain, effective cybersecurity teams can be formed. Such teams are composed of experts specific to their domain, and through cooperative work, expertise is shared to achieve shared goals. [160].

Drivers to cybersecurity:

Individual:

- Confidence. Confidence is a very useful soft skill in the cybersecurity context. It leads to better work outcomes and improved and fair interaction among colleagues. The way to increase it is to create flexible learning environments that allow for greater cooperation among colleagues and greater interactivity [154]. Trust is a key point in cybersecurity training because, as mentioned earlier, it helps people relate to and help each other against cyber attacks.
- Motivation. Two of the most important elements in the effectiveness of cybersecurity are user engagement and motivation. An employee who is committed and motivated is more inclined to team up with his colleagues leading to better overall performance [140], [141], [154].

Organizational:

- Promotes the feeling of belonging to a group. A sense of belonging to a group and organization increases awareness and contributes significantly to learning, efficient knowledge management, and problem-solving. This feeling, in the context of cybersecurity, is critically important for support in the phases pre, post, and during an attack [121].

- Cultivate organizational & team leadership skills. Successful teamwork requires cultivating an organizational culture that promotes leadership skills and peers coaching practices. Circumstances characterized by trained mentors and coaches generate more ideas and critical thinking. The ability also to be able to discuss as a team enhances a sense of responsibility and loyalty [161].
- Strengthening people and building employee engagement: Employee engagement is defined as a feeling of pride in belonging to the organization, a willingness to put up effort to ensure its success, and a sense of self-identification. The main factor influencing employee engagement is a sense of being valued and involved, which includes aspects like participation in decision-making, the degree to which employees feel free to express their opinions, the opportunities to advance in their careers, and the degree to which the company cares about the health and well-being of its workers [127]. In the context of cybersecurity, a strong sense of corporate ownership is linked to an increase in proactive ability to handle a cyber incident and an increase in loyalty and awareness toward risks.

Work:

- Show respect for the other person: All other aspects of managerial behavior were shown to be subordinate to respect. According to social psychologists, having respect for others is essential for both groups' functioning and people's well-being. It is proven that improving generalized respect relies on teamwork [162].
- Enhance team communication for overall results: Team communication and cooperation can be greatly improved by implementing direct and indirect communication channels within virtual workplaces. To establish shared thinking, shared planning, and shared understanding, a huge amount of collaborative work is necessary. Since communication is a crucial component of developing a collaborative culture, it has emerged as one of the ongoing problems that team members must overcome in order to complete any project successfully [163].
- Strong feedback system: Companies should implement a performance management system that keeps supervisors and staff members accountable for the level of involvement they have demonstrated. Employee engagement levels can be regularly surveyed to identify motivating factors. Results should prompt managers to create clear, measurable, accountable, and time-bound action plans [127]. By making the employee accountable for his or her performance will be encouraged to actively participate by being a team player.

Barriers to cybersecurity:

Individual:

- Interpersonal conflict: Interpersonal conflict refers to the representation of incompatibility, disagreement, or difference between two or more interacting individuals [129]. Conflict is inevitable. It is negative when it leads to violence, undermines the communication relationship between the parties involved in the conflict, stimulates people to become uncooperative, or prevents the parties from addressing real issues or problems [130].
- Lack of soft skills: There are four basic areas of competencies and skills that are needed by cybersecurity personnel: technical skills, non-technical (soft skills), implementation skills, and managerial skills [154]. Soft skills related to the capacity to interact effectively and harmoniously with other people are essential. If there is a high lack of soft skills, people in a company will tend not to collaborate bringing poor work results. In the cyber context, when people do not collaborate and there is a tense cyber-related risks increase.

Organizational:

- Management based on one-way communication: Managers should encourage dialogue. The workforce should feel engaged and clear and consistent communication established. A friendly environment will increase a sense of belonging and engagement in teamwork [127].
- Absence of a security-oriented organization: In teamwork, everyone must recognize their role and contribute to corporate cybersecurity [153].

Work:

- Poor performance of team members: According to the literature, teams are more likely to provide training to a member who consistently performs poorly when it is believed that he or she lacks the required skills (e.g. cyber security skills). Members are more likely to try to motivate or reject the underperformer in circumstances where they believe the person just lacks motivation. It is important in teamwork management to evaluate performance and invest in the most critical areas [164].
- Poorly managed team conflict: Disagreements among team members are common and predictable. Healthy teams discuss opposing issues and viewpoints because doing so makes their decisions stronger and more grounded. Discussion should be facilitated by the team leader or another team member. One constructive method of dialogue is to stipulate a behavioral agreement. One should avoid situations in which one of the employees feels a sense of revenge that can lead to outward leaks of information, eventually causing a data breach [165].

Table 7 - Human Factor: Teamwork

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Teamwork	*Individual*	Confidence Motivation	Interpersonal conflict Lack of soft skills
	Organizational	Promotes the feeling of belonging to a group Cultivate organizational & team leadership skills Strengthening people and building employee engagement	Management based on one-way communications Absence of a security-oriented organization
	Work	Show respect for the other person Enhance team communication for overall results Strong feedback system	Poor performance of team members Poorly managed team conflict

6. **Fatigue:**

Fatigue results from mental or physical effort or illness. One type of occupational disengagement peculiar to cybersecurity is called cybersecurity fatigue. Overexposure to cybersecurity instructions such as training or actions required by cybersecurity procedures and practices can lead to fatigue and aversion to the topic (e.g., forced updating of passwords). Interestingly, there are also two different types of cybersecurity fatigue: attitudinal fatigue (such as the belief that cybersecurity is not important) and cognitive fatigue (e.g., the habit of misbehavior) [128].

Drivers to cybersecurity:

Fatigue is a negative human feeling and factor with which no cybersecurity benefit is associated; for that reason, it cannot be considered a driver at any level (individual, organizational, and work).

Barriers to cybersecurity:

Individual:

- Personal characteristics. Fatigue is a complex biological phenomenon that is caused by several factors including personal characteristics and habits such as time awake, time of day, health, and off-duty lifestyle [159].

- Cognitive-type fatigue. The maximum amount of cognitive resources that a person can commit to security issues. In this case, for example, employee behavior could deteriorate as a result of prior effort [128].
- Attitudinal-type fatigue. It matters how the employee feels about cyber security. According to the literature, there are three main reasons why people have a bad attitude toward cyber security: they don't grasp the costs and advantages, they react, or they lack moral conviction [128].
- Tiredness. For example, an employee who is tired of being told what to do may feel exhausted and stressed due to the extreme pressure and the over-supervision within the workplace. In this sense, their fatigue is advice-related [128].

Organizational:

- Poor work planning. Work planning is useful in developing a strategy to protect an organization's system. The planning approach should include activities such as: assessing cybersecurity risk; identifying important business objectives and IT resources needed to achieve them and implementing effective communication channels. If work planning is deficient, employees would experience fatigue due to, for example, excessively long shifts, which could result in unintentional errors or misbehavior [132].
- Reactance. Tired employees rely on illogical decision-making processes. Reactivity is an example of this. This term is meant in this circumstance a negative emotional reaction caused by a sense of threat or loss of freedom on their decisions. There is a desire to recover a sense of independence, often in challenging authority. This is the case, for instance, with intentional disobedience when one feels that security regulations limit their freedom. For this reason, some researchers believe that behavioral guidelines and stricter controls at work can have the opposite effect and encourage negative employee behavior. Reactivity is considered a form of attitude-based fatigue, as it results in defiance following overly restrictive recommendations made at work [128].

Work:

- High workload. Fatigue is caused by several factors including a high workload [166].
- Technostress. Techno-overloaded employees feel that technology is adding workload rather than making it lighter generating a sense of fatigue and cybersickness. In addition, technology is often perceived as complex and constantly evolving generating in the workforce a sense of fatigue related to the perception that they will never keep up with it [128].

Table 8 - Human Factor: Fatigue

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Fatigue	*Individual*		Personal characteristics Cognitive-type fatigue Attitudinal-type fatigue Tiredness
	Organizational		Poor work planning Reactance
	Work		High workload Technostress

7. **Resources:**

   Resources are the assets necessary to produce a product or deliver a service. In the context of cybersecurity, the business architecture of a company comprises a combination of people, processes, and technology [118]:

   - People: human resources must be in adequate numbers, and they need to understand and respect basic data security principles, such as choosing complex passwords, being wary of attachments in emails, and backing up data.
   - Processes: companies should have a defined procedure to handle both attempted and successful cyber-attacks.
   - Technologies: technologies are essential to provide companies and individuals with the cybersecurity appliances they need to protect themselves from cyber-attacks. Three main entities must be protected: endpoint devices such as computers, smart devices, and routers; and networks and the cloud. Common technology used to protect these entities includes next-generation firewalls, DNS filtering, malware protection, antivirus software, and email security solutions.

   Drivers to cybersecurity:

   Organizational:

   - Automation and intelligent tools. They can fill shortfalls in knowledge and resources. Automation can help in reducing human error which brings to cyber incidents [117].
   - Adoption of frameworks. International organizations, academic institutions, corporations, and governments have been actively working to develop cybersecurity frameworks in order to provide a tool for organizations to efficiently lead and manage their resources by taking a strategic approach to

efficient cybersecurity assurance. In reality, companies may use cybersecurity frameworks to develop recommendations for the successful application of cybersecurity standards, allowing them to be better prepared to identify, detect, and respond to cyberattacks [167].

Work:

- Ensure that employees have everything they need to do their jobs. Managers are responsible for ensuring that staff members have access to all necessary information, financial, and material resources to perform their jobs [127]. An optimal work environment is positively correlated to a reduction of exposition to cyber attacks.
- Total Productive Maintenance (TPM). Maintenance is the primary duty of keeping a system working and preventing failure. It is a concept that is commonly associated with manufacturing production systems, but because of the increasing trend of adopting smart features in manufacturing systems to improve productivity, quality, and profit, the interconnectivity at the production level makes them vulnerable to cyber threats. TPM aims for 0% breakdown, slowdowns, and flaws, as well as to make the production environment safe and in perfect condition. However, cybersecurity threats may directly disrupt these goals by causing system failure, slowdowns, and quality issues, hence TPM in a cybersecurity context involves the maintenance of all computer systems and resources required to maintain an effective defense against cyber-attacks [168].

Barriers to cybersecurity:

Organizational:

- Automation and intelligent tools. Although as mentioned above, automation helps to reduce human error, excessive automation can lead to an increase in cyber threats [169]. According to "Cybersecurity, Differently," each member of the broader socio-technical system should be considered as an equal partner rather than just a substitute, rather than excluding humans from the system. The use of each partner's strengths is promoted to foster a sense of "teamwork" or synergy. Misunderstandings, security flaws, or "automation surprises" [170] among participants in an interaction can result from using automation in a way that excludes humans from the system [117].
- Lack of financial capacity. Investing in cybersecurity is expensive. The organization could adopt a low-cost approach due to its limited budget. It has been noticed that there is no objection against investing in cybersecurity, while financial responsibilities prevented from doing so. A lack of financial capacity

can prevent an organization from having all the resources it needs for efficient cybersecurity [171].

- Poor work planning. An effective approach to safeguard an organization's system can be developed with the use of work planning. A cybersecurity risk assessment should be the first step in the planning process. This assessment should identify for instance critical IT resources. The organization could not have the necessary resources to accomplish its objectives if there is poor work planning [132].
- Lack of management skills. People are the main resource for organizations. A shortage of skills leads to a reduction in those resources. Therefore, an investment must be made in programs that incentivize the enhancement of managerial and technical skills [154].

Table 9 - Human Factor: Resources

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Resources	*Individual*		
	Organizational	Automation and intelligent tools Adoption of frameworks	Automation and intelligent tools Lack of financial capacity Poor work planning Lack of management skills
	Work	Ensure that employees have everything they need to do their jobs Total Productive Maintenance	

8. **Pressure:**

Pressure is the psychological stress associated with expectations to perform well in a situation. Some pressure can actually be good for performance, research shows that an optimal amount of pressure can make you perform better. But too much pressure can lead to worsening performance and even freezing or choking [164], [172], [173].

Drivers to cybersecurity:

Individual:

- Challenge demands. Challenge demands are defined as high-pressure work-related demands or circumstances that, although potentially stressful, are related to good outcomes and potential gains for individuals [164], [172].

Barriers to cybersecurity:

Individual:

- Personal characteristic. Some people work well under pressure. However, for many others, this situation resulting from deadlines and difficult tasks can be critical. The ability to work under pressure and stress can be improved, but an individual's ability to handle it depends mainly on personal characteristics, past experiences, and personality [154], [173]
- Tiredness. A sense of tiredness can be associated with those situations in which the worker is constantly under pressure and feels excessive supervision and regulation upon himself/herself [128].
- Poor work planning. As anticipated in the previous paragraphs, if there is poor work planning, employees could experience a lot of pressure and stress due to, for example, a complicated task that doesn't match their abilities or an unfair and unbalanced distribution of the workload [132].
- Complexity of the norms. Pressure and stress can result from the complexity of regulations. Security policies and procedures are often complex. The technical language and amount of information can be difficult to understand, requiring employees to spend time and effort to learn them [174].

Work:

- Work pressure. Work pressure refers to pressure resulting from high workloads, or work demands from the organization that does not meet the employee's abilities [154].
- Time pressure. Time pressure refers to *"objective or subjective perceived limitation of the available time needed to consider information or to take a decision"* [175]. Time pressure can be objective, since we may have time constraints resulting from explicit deadlines. It can be subjective in cases where we feel pressured by requests for tasks to be done with urgency and need to process many tasks and information [176][177]. Users in these situations perceive a lack of their effectiveness in meeting IT security requirements. For example, developers perceive limited time as an obstacle to their ability to code securely [154].

Table 10 - Human Factor: Pressure

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Pressure	*Individual*	Challenge Demands	Personal characteristics Tiredness
	Organizational		Poor work planning Complexity of the norms
	Work		Work Pressure Time Pressure

9. **Assertiveness:**

Assertiveness is the ability to express one's viewpoint, opinions, ideas, or rights without undermining those of others. It is essential in communication between employees and the employer. A deficiency in this skill can impair employees' job performance. Assertiveness is a fundamental behavior for creating and maintaining positive relationships at work and facilitating team functioning and decision-making in critical situations.

Drivers to cybersecurity:

Individual:

- Personal characteristics. In a study conducted by [178], it was found that people who scored high on assertiveness were likely to be more determined, competitive, energetic, and work with greater drive and purpose. Assertiveness was found to be significantly related to work engagement [178]. This may explain why cybersecurity professionals scored higher than regular IT employees. Their tasks have a greater impact on the organization and people by ensuring their security and indirectly their safety. They are also required to keep up to date with new technologies which lead to greater motivation.

Organizational:

- Promotes the feeling of belonging to a group. Assertiveness allows the employee to ask questions, express displeasure, and contribute ideas, which gives a sense of belonging [179]. If the employee has a sense of being part of a group, he or she will be able to express his or her emotions and assert his or her thoughts without prevaricating those of others, contributing to better work performance.

Barriers to cybersecurity:

Individual:

- Fear. Fear is an individual barrier to assertiveness in the context of cybersecurity. “Fear” to communicate a mistake to one of our superiors for the negative consequences, or to communicate with one of our colleagues because we have “fear” of their judgment [117]. If an employee is afraid, he will be more reluctant to assert his thoughts, and this will lead to ineffective collaboration.

Work:

- Lack of initiative. People with low assertiveness tend to let others control the group, showing a lack of initiative [180]. In the context of cybersecurity, a lack of initiative can result in the inability to take responsibility or make an important decision to resolve a cyber attack.

Table 11 - Human Factor: Assertiveness

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Assertiveness	*Individual*	Personal characteristics	Fear
	Organizational	Promotes the feeling of belonging to a group	
	Work		Lack of initiative

10. **Stress:**

Stress is defined as a condition that causes a psychophysiological response in an individual that deviates from a state of equilibrium [181]. In the cybersecurity context, stress elements can have various sources, such as management efforts, task and role assignments, interpersonal conflicts, and procedures and rules.

Drivers to cybersecurity:

Individual:

- Challenge demands. As in the case mentioned for pressure, challenge demands are defined as high-pressure work-related demands or circumstances that, although potentially stressful, are related to good outcomes and potential gains for individuals [164], [172].

Barriers to cybersecurity:

Individual:

- Personal characteristics. While for some people working under pressure and stress could result in better performance for others the high stress coming from, deadlines and difficult tasks can create extreme discomfort. The ability to work under pressure and stress can be improved but our ability to handle it mainly on personal characteristics, past experiences, and personality [154], [173].

- Tiredness. A sense of fatigue may be associated with stressful situations in which the worker feels high responsibilities and a large amount of work on him or her [128].
- Interpersonal conflict. As mentioned earlier, interpersonal conflict refers to the representation of incompatibility, disagreement, or difference between two or more interacting individuals [129]. In the presence of interpersonal conflict, the employee tends to experience stress in the workplace [130].

Organizational:

- Poor work planning. As anticipated in the previous paragraphs, poor planning generates in employees an increase in stress related to a lack of detailed information, clear work schedules, and division of tasks [132].
- The complexity of the norms. Pressure and stress could result from the complexity of the norms. In addition, cyber security issues tend to cause a lot of stress as the worker perceives how a misunderstanding of procedures or rules can cause a major incident for the organization [174].
- Lack of resources. According to researchers, stress could result from the shortage of resources. Organizational leaders need to ensure an adequate balance between the workplace environment and the employee [182], [183].

Work:

- Work pressure. Work pressure refers to pressure resulting from high workloads and is strongly correlated with increased stress. [175].
- Time pressure. As in the previous case, the time pressure discussed above also appears to be correlated with increased stress in workers [154].
- Technostress.: New or unclear technology might cause employee anxiety and a bad attitude toward the technology. [184] refers to these technologically based pressures as "technostress". According to [174], the three elements that contribute to technological stress and are essential to cyber security are overload, complexity, and ambiguity of technology.

Table 12 - Human Factor: Stress

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Stress	*Individual*	Challenge Demands	Personal characteristics. Tiredness Interpersonal conflict
	Organizational		Poor work planning. Complexity of the norms Lack of resources
	Work		Work Pressure Time Pressure Technostress

11. **Awareness:**

Awareness is the *"knowledge that something exists, or understanding of a situation or subject at the present time based on information or experience"*[185]. In the context of cybersecurity, it can be defined as a learning process that lays the foundation for training, changing individual and organizational attitudes to realize the importance of security and the negative consequences in the event of an incident. Many studies state how this is achieved only through training and not only through awareness-raising. In awareness-raising activities, the learner is the recipient of the information, whereas in a training environment he or she takes a more active role [60].

Drivers to cybersecurity:

Individual:

- Success stories. Publishing success stories and praising employees who detect attacks are all useful practices to alert other employees to improve cybersecurity [60]. This practice raises awareness among employees, increasing the ability to identify attacks.
- Motivation. Education and training enhance user engagement and motivation. An employee who is committed and motivated is more inclined to increase his or her knowledge and awareness [140], [141], [154].

Organizational:

- Awareness campaigns. Awareness campaigns are critical to improving cybersecurity. When these are done carelessly or performed in a repetitive and monotonous manner, the opposite effect is achieved. Employees will come across as disinterested in cybersecurity practices. When instead organized and attractive, they contribute to efficient cybersecurity by making people in the organization aware of the risks involved [103].

- Training. The topic of training related to increasing awareness of cyber risks is highly stressed in the literature. Workers if properly trained can be the primary driver for more effective cyber security [113], [151]. Interesting are the studies that emphasize the importance of innovative approaches (e.g., Virtual Reality, gaming, chatbots) to training that are found to be more effective in increasing cyber security awareness [186]–[188].
- Cyber Hygiene (CH). Cyber Hygiene practices introduced before collaborate to increase organizational cybersecurity culture and user awareness. [150].

Work:

- Cyberdrill. This is a training process that simulates a cyber-attack on employees or people whose work is related to cyber-incident response with the goal to improve skills and raise awareness [151].
- Strong Feedback system. A structured and effective feedback system helps increase employee awareness. They will feel supported and followed in an ongoing training process [127].

Barriers to cybersecurity:

Individual:

- Cognitive-type fatigue. The maximum amount of cognitive resources that a person can commit to security issues. In this case, for example, employee behavior could deteriorate as a result of prior inefficient awareness campaigns [128].
- Tiredness. For example, an employee who is tired of participating in an awareness campaign or training may feel exhausted and stressed due to the extreme pressure and the over-supervision within the workplace [128].
- Attitude toward risk. A person with a higher risk propensity might be less afraid of the consequences of a cyber-attack and consequently not perceive the severity of not being aware of cybersecurity issues [147].
- Overtrust in cybersecurity devices. Overtrust is a barrier to cybersecurity awareness. Employees do not always have a clear understanding of how cybersecurity systems work. This leads them to be overconfident and unaware of the risks [136].

Organizational:

- Lack of investment. According to the Security Awareness Report by SANS of 2022 [189], most organizations does not invest enough in training and awareness campaign.
- Lack of staff. According to [189], most organizations don't have enough staff dedicated to awareness programs. This problem affects not only small companies but also larger companies where most of the resources responsible for awareness programs, are also employed in other tasks and areas not devoting full concentration to these campaigns.
- Absence of a security-oriented organization. An organization that does not invest in security causes employees to become disinterested in the issue and unconcerned about it. This does not allow employees to generate cybersecurity awareness [153].

Work:

- Lack of time. According to [189], both campaign managers and employees devote a very low percentage of their work to training because they are busy with other tasks.

Table 13 - Human Factor: Awareness

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Awareness	***Individual***	Success stories Motivation	Cognitive- type of fatigue Tiredness Attitude toward risk Overtrust in cybersecurity devices
	Organizational	Awareness campaigns Training Cyber Hygiene	Lack of investments Lack of staff Absence of a security-oriented organization
	Work	Cyberdrill Strong feedback system	Lack of time

12. **Norms:**

Among the most widely accepted definition of norms is the one given by [149] which define norms as "*collective expectation for the proper behavior of actors with a given identity.*" The development of norms requires a shared belief about proper behavior for actors in a community (e.g. organization).

Drivers to cybersecurity:

Individual:

- Motivation. An employee who is committed and motivated is more inclined to follow good cybersecurity behavior and to follow desirable security norms [140], [141], [154].
- Employee satisfaction. An employee who is satisfied with his position and his job is more inclined to stay loyal to the company and to be strict in following the organization's security policies [144].

Organizational:

- Awareness campaigns. Cyber awareness campaigns help the organization convey to employees not only an awareness of risks but also the rationale behind imposed rules, procedures, and practices [190].
- Training. Training campaigns are critical to successfully passing concepts related to cybersecurity norms. The topic turns out to be complex and tedious, it is necessary to invest in innovative training that manages to reach employees with different backgrounds.

Barriers to cybersecurity:

Individual:

- Cognitive-type fatigue. Workers may experience a type of cognitive fatigue related to a large number of norms to remember and comply with [128].
- Attitude toward risk. A person with a higher risk propensity might have less fear of the consequences related to noncompliance with organizational norms [147].

Organizational:

- Complexity of the norms. The complex and technical language of standards can be very difficult to understand and requires employees' time and effort Absence of a security-oriented organization: The absence of a corporate cybersecurity culture leads the organization to underestimate cybersecurity norms, policies, and procedures [153].

Table 14 - Human Factor: Norms

Human Factors	*PIFs*	*Drivers to cybersecurity*	*Barriers to cybersecurity*
Norms	*Individual*	Motivation Employee satisfaction	Cognitive-type fatigue Attitude toward risk
	Organizational	Awareness Campaigns Training	Complexity of the norms
	Work		

4.4. NIST & Human Factors

In this section, the research will now focus on the third question mentioned above: *which role do the defined human factors play in the functions and categories of the NIST framework?* The research will investigate the role of Dupont's human factor concerning the functions and categories of the NIST framework [60]. For a detailed explanation of the NIST framework, the reader is invited to refer to section 2.4.2.

The NIST framework intends to equip organizations and enterprises with optimal risk management practices that may be adopted to increase critical infrastructure security and resilience. The National Institute of Standards and Technology (NIST) considers risk management as an iterative process of risk identification, risk assessment, and risk reduction.

While the NIST framework gives a nicely ordered account of enterprises' and organizations' cybersecurity activities, it fails to convey the idea that humans are part of the system and inherent risk. To move beyond the current risk framework promulgated by NIST, the risk assessment needs to be more holistic and incorporate humans and related risk factors into a single model [94].

In particular, the framework fails to emphasize the idea that humans, whether users, defenders, or attackers, can introduce risks into the network. This is the case for untrained or aware users. Even defenders who are less skilled, or tired, or if they are within threats can introduce risk. Furthermore, the framework lacks the vision in which humans can also mitigate risk in a cyber resilience perspective.

Defenders can put in place basic protections and then monitor intrusions on the system to see whether protections have been violated and what needs to be done to counteract malware and fix system damage. Users can mitigate risk by being aware of spam and phishing efforts and ensuring that their own system resources are suitably secured. As a result, human-dependent parameters must be incorporated into a comprehensive assessment of cybersecurity threats.

Our conceptual solution proposes to take the NIST framework as a basis and include in it the human dimension by integrating HF and cybersecurity situations that have emerged in the literature.

Two representations are provided below. A first representation that is more extensive involves the inclusion of an additional "human factors" column to the already extensive basic framework. The human factors reported in the column are those found to be most related in terms of their possible impact on that particular NIST category. The second representation (Figure 13) shows a summary of the factors involved in the risk assessment pathway proposed by NIST.

Starting from left to right, first, the NIST function (e.g., IDENTIFY) is defined by the underlying which is the main strategic objective. Next, it is mentioned the category, indicating the area of expertise to which the framework is referring (e.g., Asset Management (ID.AM)). Finally, before detailing the information into subcategories and informative references, containing both strategic and technical best practices, this research new framework incorporates a Human Factors column. This column helps identify which are the possible human factors that could negatively affect the Category and stress the importance to consider best practices related to them to reach the category objective.

For example, in the IDENTIFY function, aspects useful for managing cybersecurity risk to systems, people, assets, data, and capabilities are analyzed. Its first category Asset Management stresses the importance of data, personnel, devices, systems, and facilities that enable the organization to achieve business purposes. In this situation, it is recommended that the organization inventory all devices, assign roles and responsibilities, and map organizational communication and data flows. However, the new column draws attention to practices that are useful in achieving the ultimate objective. It points out, for example, how if there is no proper distribution of resources and training of skills (*knowledge*), employees might feel a sensation of *Stress*. They are supposed to deal with a complex scenario without having the necessary resources to perform a given task.

1. IDENTIFY

Develop an organizational understanding to manage cybersecurity risk to systems, people, assets, data, and capabilities.

Asset Management (ID.AM):

"The data, personnel, devices, systems, and facilities that enable the organization to achieve business purposes are identified and managed consistent with their relative importance to organizational objectives and the organization's risk strategy."

Human Factors:

- Communication. Responsibilities within teams are assigned and official communication channels are established. To ensure optimal human asset management, establishing efficient internal communication is the key. Employees must feel part of a group and have the feeling that they can communicate with management without repercussions. In addition, clarification of roles and responsibilities is essential to avoid interpersonal conflicts and facilitate better work planning [121], [122].
- Resources. Management ensures that employees have all the equipment (both human e technical) they need to perform their duties and that it is in perfect condition [127], [168]. In the context of cybersecurity, the business architecture of a company comprises a combination of people, processes, and technology [118]. The organization must adopt detailed planning of both the budget, to ensure that it has the adequate financial capacity to fund all assets, and of the work, to ensure that all resources are available in the right place at the right time.
- Knowledge. If the organization does not know how to achieve its objectives, it is not possible to manage resources correctly. To best optimize assets, the organization must have a 360-degree understanding of both the resources at its disposal and the strategic goals it wants to achieve. This means being able to prioritize them based on their criticality and their business value (Al-Dawod et al.,2021).
- Stress. If the employee does not have the proper materials available or the responsibilities within his or her team are unclear, he or she may experience stress. According to researchers, stress could result from the shortage of resources. Organizational leaders need to ensure an adequate balance between the workplace environment and the employee [182], [183], [191]. In the cyber environment where even with all the resources at hand it is difficult to make mistakes not having them or having them inadequate equates to a very high probability of errors for employees and this can be stressful.

Business Environment (ID.BE):
"The organization's mission, objectives, stakeholders, and activities are understood and prioritized; this information is used to inform cybersecurity roles, responsibilities, and risk management decisions."

Human Factors:

- Communication. Strategic and non-strategic goals of the organization are communicated promptly and roles and responsibilities are assigned. Creating a functional business environment, all starts with internal communication. The information must be shared quickly and correctly, and responsibilities and roles clear and well-defined. Activities are communicated promptly and clearly to avoid poor work planning problems [122], [132].

- Stress. If the responsibilities within his or her team are unclear or the assigned tasks do not meet the employee's skills, he or she may experience stress. During the assignment of the various tasks, the organization must consider the personal characteristics of the employees such as their personality and experience to ensure that everyone can work at their best. In fact, stress can arise from the conflict between two members or the unsized workload. In creating teams and assigning activities, management must ensure that team members are compatible, and that the activities meet the skills of the employees [122], [130], [159].
- Fatigue. If roles and responsibilities are not clear or the work schedule is not properly balanced the employee may experience cybersecurity fatigue [128], [183].
- Awareness. Organizations depend critically on top management's ability to create a business environment in which everyone is aware of cybersecurity issues. A good way is to organize awareness campaigns but to get employees to participate thoughtfully, organizations must create an environment in which the employee is motivated and stimulated to take care of his/her organization. The top management has a significant role in the influence of organizational culture and knowledge and thereby the risk awareness of a company, which in turn has an impact on their cybersecurity. It is the responsibility of the top management to delegate tasks that enhance risk awareness [140], [141], [154].
- Knowledge. To achieve the goals it has set for itself the organization should have the knowledge to make the right decisions. From assigning optimal roles to selecting activities, nothing can be chosen correctly without knowledge of the environment in which one works. The most effective way to increase knowledge is through training, but it is essential to make it challenging and nonrepetitive to ensure the full attention and participation of those involved. Practically, having better knowledge, give improved relations with stakeholders and, thereby company survival as a result [60], [140], [141], [154].

Governance (ID.GV):

"The policies, procedures, and processes to manage and monitor the organization's regulatory, legal, risk, environmental, and operational requirements are understood, and the management of cybersecurity risk is informed."

Human Factors:

- Norms. Corporate norms are communicated efficiently. While considering the importance of safety, they are easy to implement and do not require exaggerated effort in either understanding or implementation [127].
- Stress. If the norms are too complex the employee could experience technostress [128].
- Fatigue. The employee may become cognitively fatigued if the norms are very complicated or repetitious [128].
- Pressure. If the norms are too complex and difficult to implement, due to too much pressure the employee could make mistakes [174].

Risk Assessment (ID.RA):

"The organization understands the cybersecurity risk to organizational operations (including mission, functions, image, or reputation), organizational assets, and individuals."

Human Factors:

- Awareness. Everyone in the organization is aware of cyber risks related to work operations and attacks on personal devices. For instance, the organizational members need to be aware of digital activities, like downloading software or any kind of information disclosure to third parties. This is because it can reveal confidential information about the company. However, security awareness is a perpetual learning process that takes place on an organizational level, where all individuals in the organization are included and some possess certain duties to manage it. This results in a need for the management to comprehend and design strategies to handle and improve cybersecurity [192].
- Complacency. If someone within the organization is unaware of cyber risks or lacks the knowledge to handle the devices at their disposal they may feel a sense of complacency, convinced that they have the situation under control. People's seeming complacency is mostly due to the lack of a significant and damaging cyber incident that has disrupted a critical service in their lifetime or to their overtrust in cybersecurity devices. Employees often do not have a broad understanding of how devices devoted to cybersecurity work. This leads them to overtrust them and to become distracted, confident that the devices can form an effective line of defense autonomously, without their supervision [94], [134].

Risk Management Strategy (ID.RM):

"The organization's priorities, constraints, risk tolerances, and assumptions are established and used to support operational risk decisions."

Human Factors:

- Knowledge. to understand the cybersecurity risks and adopt the right Risk Management Strategy the organization, need to have a deep knowledge of the best cybersecurity practices. Asking an employee to manage the company's cybersecurity without possessing the knowledge and tools necessary to perform the task will result in poor-quality cybersecurity. Management needs to know the company's digital network, including how systems are integrated, the inflow and outflow of information, and frequently used digital pipes, and disseminate this knowledge. By doing so, it becomes easier to discover atypical activity and thus investigate and even identify potential hackers [192].
- Stress. If the processes are too complex employees may experience significant stress as a result of a difficult assignment that is above their ability. Work planning can help a lot to develop a great strategy to protect the system of an organization. The planning approach ought, to begin with a cybersecurity risk assessment and then the organization needs to establish crucial IT resources, implement effective communication channels and identify prospective cyberattacks as well as how probable the attacks are to happen and what kind of business impact they could have. If there is poor work planning, employees could experience a lot of pressure and stress due to, for example, a complicated task that doesn't match their abilities [132].

Supply Chain Risk Management (ID.SC):
"The organization's priorities, constraints, risk tolerances, and assumptions are established and used to support risk decisions associated with managing supply chain risk. The organization has established and implemented the processes to identify, assess and manage supply chain risks."

Human Factors:

- Norms. To establish a protected supply chain, pre-established standards should be adopted jointly with partners. A lack of accepted standards and guidelines is hindering the development of robust cyber defenses. [193], [194] argue that supply chain partners must be more transparent with each other on security and should combine security resources and know-how to deal with increasingly sophisticated cyber risks. [195] recommend that supply chain integration, by aligning systems and processes, will yield better returns through standardized ways of working, shared security objectives, and better general communication.
- Awareness. Supply chain members should also be aware of their partner's cyber risks to foster safer collaboration. The propagation of cyber consequences means companies cannot afford to focus only on their security systems and must also be aware of their partner's security conditions [196]. Organizational

security system mitigates cyber-attack by securing physical assets, adhering to set guidelines, and by raising awareness among employees. Information sharing, collaborative risk management, and adaptability are found to be key strategies for supply chain security [166].

Table 15 - NIST function IDENTIFY & Human Factors

Function	Category	Human Factors	Subcategory	Informative References
IDENTIFY (ID)	**Asset Management (ID.AM):** The data, personnel, devices, systems, and facilities that enable the organization to achieve business purposes are identified and managed consistent with their relative importance to organizational objectives and the organization's risk strategy.	**Communication:** Responsibilities within teams are assigned and official communication channels are established **Resources:** Management ensures that employees have all the equipment (both human e technical) they need to perform their duties and that it is in perfect condition **Knowledge:** If the organization does not know how to achieve its objectives, it is not possible to manage resources correctly **Stress:** If the employee does not have the proper materials available or the responsibilities within his or her team are unclear, he or she may experience stress	**ID.AM-1:** Physical devices and systems within the organization are inventoried	**CIS CSC** 1 **COBIT 5** BAI09.01, BAI09.02 **ISA 62443-2-1:2009** 4.2.3.4 **ISA 62443-3-3:2013** SR 7.8 **ISO/IEC 27001:2013** A.8.1.1, A.8.1.2 **NIST SP 800-53 Rev. 4** CM-8, PM-5
			ID.AM-2: Software platforms and applications within the organization are inventoried	**CIS CSC** 2 **COBIT 5** BAI09.01, BAI09.02, BAI09.05 **ISA 62443-2-1:2009** 4.2.3.4 **ISA 62443-3-3:2013** SR 7.8 **ISO/IEC 27001:2013** A.8.1.1, A.8.1.2, A.12.5.1 **NIST SP 800-53 Rev. 4** CM-8, PM-5
			ID.AM-3: Organizational communication and data flows are mapped	**CIS CSC** 12 **COBIT 5** DSS05.02 **ISA 62443-2-1:2009** 4.2.3.4 **ISO/IEC 27001:2013** A.13.2.1, A.13.2.2 **NIST SP 800-53 Rev. 4** AC-4, CA-3, CA-9, PL-8
			ID.AM-4: External information systems are catalogued	**CIS CSC** 12 **COBIT 5** APO02.02, APO10.04, DSS01.02 **ISO/IEC 27001:2013** A.11.2.6 **NIST SP 800-53 Rev. 4** AC-20, SA-9
			ID.AM-5: Resources (e.g., hardware, devices, data, time, personnel, and software) are prioritized based on their classification, criticality, and business value	**CIS CSC** 13, 14 **COBIT 5** APO03.03, APO03.04, APO12.01, BAI04.02, BAI09.02 **ISA 62443-2-1:2009** 4.2.3.6 **ISO/IEC 27001:2013** A.8.2.1 **NIST SP 800-53 Rev. 4** CP-2, RA-2, SA-14, SC-6
			ID.AM-6: Cybersecurity roles and responsibilities for the entire workforce and third-party stakeholders are established	**ISA 62443-2-1:2009** 4.3.2.3.3 **ISO/IEC 27001:2013** A.6.1.1 **NIST SP 800-53 Rev. 4** CP-2, PS-7, PM-11

Function	Category	Human Factors	Subcategory	Informative References
IDENTIFY (ID)	**Business Environment (ID.BE):** The organization's mission, objectives, stakeholders, and activities are understood and prioritized; this information is used to inform cybersecurity roles, responsibilities, and risk management decisions.	**Communication:** Strategic and non-strategic goals of the organization are communicated promptly and roles and responsibilities are assigned **Stress:** If responsibilities within teams are unclear or the assigned tasks do not meet the employee's skills, he or she may experience stress. **Fatigue:** If roles and responsibilities are not clear or the work schedule is not balanced the employee may experience cybersecurity fatigue **Awareness:** Organizations depends critically on top management's ability to create a business environment in which everyone is aware of cybersecurity issues **Knowledge:** The organization should have the knowledge to make the right decisions.	**ID.BE-1:** The organization's role in the supply chain is identified and communicated	**COBIT 5** APO08.01, APO08.04, APO08.05, APO10.03, APO10.04, APO10.05 **ISO/IEC 27001:2013** A.15.1.1, A.15.1.2, A.15.1.3, A.15.2.1, A.15.2.2 **NIST SP 800-53 Rev. 4** CP-2, SA-12
			ID.BE-2: The organization's place in critical infrastructure and its industry sector is identified and communicated	**COBIT 5** APO02.06, APO03.01 **ISO/IEC 27001:2013** Clause 4.1 **NIST SP 800-53 Rev. 4** PM-8
			ID.BE-3: Priorities for organizational mission, objectives, and activities are established and communicated	**COBIT 5** APO02.01, APO02.06, APO03.01 **ISA 62443-2-1:2009** 4.2.2.1, 4.2.3.6 **NIST SP 800-53 Rev. 4** PM-11, SA-14
			ID.BE-4: Dependencies and critical functions for delivery of critical services are established	**COBIT 5** APO10.01, BAI04.02, BAI09.02 **ISO/IEC 27001:2013** A.11.2.2, A.11.2.3, A.12.1.3 **NIST SP 800-53 Rev. 4** CP-8, PE-9, PE-11, PM-8, SA-14
			ID.BE-5: Resilience requirements to support delivery of critical services are established for all operating states (e.g. under duress/attack, during recovery, normal operations)	**COBIT 5** BAI03.02, DSS04.02 **ISO/IEC 27001:2013** A.11.1.4, A.17.1.1, A.17.1.2, A.17.2.1 **NIST SP 800-53 Rev. 4** CP-2, CP-11, SA-13, SA- 14

Function	Category	Human Factors	Subcategory	Informative References
IDENTIFY (ID)	**Governance (ID.GV):** The policies, procedures, and processes to manage and monitor the organization's regulatory, legal, risk, environmental, and operational requirements are understood and the management of cybersecurity risk is informed.	**Norms:** Corporate norms are communicated efficiently. While considering the importance of safety, they are easy to implement and do not require exaggerated effort in either understanding or implementation **Fatigue:** The employee may become cognitively fatigued if the norms are very complicated or repetitious **Pressure:** If the norms are too complex and difficult to implement, due to too much pressure the employee could make mistakes **Stress:** If the norms are too complex the employee could experience technostress	**ID.GV-1:** Organizational cybersecurity policy is established and communicated	**CIS CSC** 19 **COBIT 5** APO01.03, APO13.01, EDM01.01, EDM01.02 **ISA 62443-2-1:2009** 4.3.2.6 **ISO/IEC 27001:2013** A.5.1.1 **NIST SP 800-53 Rev. 4** -1 controls from all security control families
			ID.GV-2: Cybersecurity roles and responsibilities are coordinated and aligned with internal roles and external partners	**CIS CSC** 19 **COBIT 5** APO01.02, APO10.03, APO13.02, DSS05.04 **ISA 62443-2-1:2009** 4.3.2.3.3 **ISO/IEC 27001:2013** A.6.1.1, A.7.2.1, A.15.1.1 **NIST SP 800-53 Rev. 4** PS-7, PM-1, PM-2
			ID.GV-3: Legal and regulatory requirements regarding cybersecurity, including privacy and civil liberties obligations, are understood and managed	**CIS CSC** 19 **COBIT 5** BAI02.01, MEA03.01, MEA03.04 **ISA 62443-2-1:2009** 4.4.3.7 **ISO/IEC 27001:2013** A.18.1.1, A.18.1.2, A.18.1.3, A.18.1.4, A.18.1.5 **NIST SP 800-53 Rev. 4** -1 controls from all security control families
			ID.GV-4: Governance and risk management processes address cybersecurity risks	**COBIT 5** EDM03.02, APO12.02, APO12.05, DSS04.02 **ISA 62443-2-1:2009** 4.2.3.1, 4.2.3.3, 4.2.3.8, 4.2.3.9, 4.2.3.11, 4.3.2.4.3, 4.3.2.6.3 **ISO/IEC 27001:2013** Clause 6 **NIST SP 800-53 Rev. 4** SA-2, PM-3, PM-7, PM- 9, PM-10, PM-11

Function	Category	Human Factors	Subcategory	Informative References
IDENTIFY (ID)	**Risk Assessment (ID.RA):** The organization understands the cybersecurity risk to organizational operations (including mission, functions, image, or reputation), organizational assets, and individuals.	**Awareness:** Everyone in the organization is aware of cyber risks both related to work operations and attacks on personal devices **Complacency:** If someone within the organization is unaware of cyber risks or lacks the knowledge to handle the devices at their disposal they may feel a sense of complacency, convinced that they have the situation under control	**ID.RA-1:** Asset vulnerabilities are identified and documented	**CIS CSC** 4 **COBIT 5** APO12.01, APO12.02, APO12.03, APO12.04, DSS05.01, DSS05.02 **ISA 62443-2-1:2009** 4.2.3, 4.2.3.7, 4.2.3.9, 4.2.3.12 **ISO/IEC 27001:2013** A.12.6.1, A.18.2.3 **NIST SP 800-53 Rev. 4** CA-2, CA-7, CA-8, RA- 3, RA-5, SA-5, SA-11, SI-2, SI-4, SI-5
			ID.RA-2: Cyber threat intelligence is received from information sharing forums and sources	**CIS CSC** 4 **COBIT 5** BAI08.01 **ISA 62443-2-1:2009** 4.2.3, 4.2.3.9, 4.2.3.12 **ISO/IEC 27001:2013** A.6.1.4 **NIST SP 800-53 Rev. 4** SI-5, PM-15, PM-16
			ID.RA-3: Threats, both internal and external, are identified and documented	**CIS CSC** 4 **COBIT 5** APO12.01, APO12.02, APO12.03, APO12.04 **ISA 62443-2-1:2009** 4.2.3, 4.2.3.9, 4.2.3.12 **ISO/IEC 27001:2013** Clause 6.1.2 **NIST SP 800-53 Rev. 4** RA-3, SI-5, PM-12, PM- 16
			ID.RA-4: Potential business impacts and likelihoods are identified	**CIS CSC** 4 **COBIT 5** DSS04.02 **ISA 62443-2-1:2009** 4.2.3, 4.2.3.9, 4.2.3.12 **ISO/IEC 27001:2013** A.16.1.6, Clause 6.1.2 **NIST SP 800-53 Rev. 4** RA-2, RA-3, SA-14, PM- 9, PM-11
			ID.RA-5: Threats, vulnerabilities, likelihoods, and impacts are used to determine risk	**CIS CSC** 4 **COBIT 5** APO12.02 **ISO/IEC 27001:2013** .12.6.1 **NIST SP 800-53 Rev. 4** RA-2, RA-3, PM-16

Function	Category	Human Factors	Subcategory	Informative References
IDENTIFY (ID)			**ID.RA-6:** Risk responses are identified and prioritized	**CIS CSC** 4 **COBIT 5** APO12.05, APO13.02 **ISO/IEC 27001:2013** Clause 6.1.3 **NIST SP 800-53 Rev. 4** PM-4, PM-9
	Risk Management Strategy (ID.RM): The organization's priorities, constraints, risk tolerances, and assumptions are established and used to support operational risk decisions.	**Knowledge:** to understand the cybersecurity risks and adopt the right Risk Management Strategy the organization, need to have a deep knowledge of the best cybersecurity practices **Stress:** If the processes are too complex employees may experience significant stress as a result of a difficult assignment that is above their ability	**ID.RM-1:** Risk management processes are established, managed, and agreed to by organizational stakeholders	**CIS CSC** 4 **COBIT 5** APO12.04, APO12.05, APO13.02, BAI02.03, BAI04.02 **ISA 62443-2-1:2009** 4.3.4.2 **ISO/IEC 27001:2013** Clause 6.1.3, Clause 8.3, Clause 9.3 **NIST SP 800-53 Rev. 4** PM-9
			ID.RM-2: Organizational risk tolerance is determined and clearly expressed	**COBIT 5** APO12.06 **ISA 62443-2-1:2009** 4.3.2.6.5 **ISO/IEC 27001:2013** Clause 6.1.3, Clause 8.3 **NIST SP 800-53 Rev. 4** PM-9
			ID.RM-3: The organization's determination of risk tolerance is informed by its role in critical infrastructure and sector specific risk analysis	**COBIT 5** APO12.02 **ISO/IEC 27001:2013** Clause 6.1.3, Clause 8.3 **NIST SP 800-53 Rev. 4** SA-14, PM-8, PM-9, PM- 11
	Supply Chain Risk Management (ID.SC): The organization's priorities, constraints, risk tolerances, and assumptions are established and used to support risk	**Norms:** To establish a protected supply chain, pre-established standards should be adopted jointly with partners. A lack of accepted standards and guidelines is hindering the development of robust cyber defenses.	**ID.SC-1:** Cyber supply chain risk management processes are identified, established, assessed, managed, and agreed to by organizational stakeholders	**CIS CSC** 4 **COBIT 5** APO10.01, APO10.04, APO12.04, APO12.05, APO13.02, BAI01.03, BAI02.03, BAI04.02 **ISA 62443-2-1:2009** 4.3.4.2 **ISO/IEC 27001:2013** A.15.1.1, A.15.1.2, A.15.1.3, A.15.2.1, A.15.2.2 **NIST SP 800-53 Rev. 4** SA-9, SA-12, PM-9

Function	Category	Human Factors	Subcategory	Informative References
IDENTIFY (ID)	decisions associated with managing supply chain risk. The organization has established and implemented the processes to identify, assess and manage supply chain risks.	**Awareness:** Supply chain members must also be made aware of their partners' cyber risks to foster safer collaboration	**ID.SC-2:** Suppliers and third-party partners of information systems, components, and services are identified, prioritized, and assessed using a cyber supply chain risk assessment process	**COBIT 5** APO10.01, APO10.02, APO10.04, APO10.05, APO12.01, APO12.02, APO12.03, APO12.04, APO12.05, APO12.06, APO13.02, BAI02.03 **ISA 62443-2-1:2009** 4.2.3.1, 4.2.3.2, 4.2.3.3, 4.2.3.4, 4.2.3.6, 4.2.3.8, 4.2.3.9, 4.2.3.10, 4.2.3.12, 4.2.3.13, 4.2.3.14 **ISO/IEC 27001:2013** A.15.2.1, A.15.2.2 **NIST SP 800-53 Rev. 4** RA-2, RA-3, SA-12, SA- 14, SA-15, PM-9
			ID.SC-3: Contracts with suppliers and third-party partners are used to implement appropriate measures designed to meet the objectives of an organization's cybersecurity program and Cyber Supply Chain Risk Management Plan.	**COBIT 5** APO10.01, APO10.02, APO10.03, APO10.04, APO10.05 **ISA 62443-2-1:2009** 4.3.2.6.4, 4.3.2.6.7 **ISO/IEC 27001:2013** A.15.1.1, A.15.1.2, A.15.1.3 **NIST SP 800-53 Rev. 4** SA-9, SA-11, SA-12, PM- 9
			ID.SC-4: Suppliers and third-party partners are routinely assessed using audits, test results, or other forms of evaluations to confirm they are meeting their contractual obligations.	**COBIT 5** APO10.01, APO10.03, APO10.04, APO10.05, MEA01.01, MEA01.02, MEA01.03, MEA01.04, MEA01.05 **ISA 62443-2-1:2009** 4.3.2.6.7 **ISA 62443-3-3:2013** SR 6.1 **ISO/IEC 27001:2013** A.15.2.1, A.15.2.2
			ID.SC-5: Response and recovery planning and testing are conducted with suppliers and third-party providers	**NIST SP 800-53 Rev. 4** AU-2, AU-6, AU-12, AU- 16, PS-7, SA-9, SA-12 **CIS CSC** 19, 20 **COBIT** 5 DSS04.04 **ISA 62443-2-1:2009** 4.3.2.5.7, 4.3.4.5.11 **ISA 62443-3-3:2013** SR 2.8, SR 3.3, SR.6.1, SR 7.3, SR 7.4 **ISO/IEC 27001:2013** .17.1.3 **NIST** SP 800-53 Rev. 4 CP-2, CP-4, IR-3, IR-4, IR-6, IR-8, IR-9

2. PROTECT

Develop and implement appropriate safeguards to ensure delivery of critical services.

Identity Management, Authentication and Access Control (PR.AC):

"Access to physical and logical assets and associated facilities is limited to authorized users, processes, and devices, and is managed consistent with the assessed risk of unauthorized access to authorized activities and transactions."

Human Factors:

- Fatigue. A condition of tiredness and fatigue due to the continuous validation of one's identity, accompanied by the requirement to create complicated passwords, passphrase deadlines, and additional credentials to gain access, is referred to as authentication fatigue [140]. Cybersecurity professionals must model and manage access that is easy to use and less frustrating [197].
- Norms. The organization must give its employees various authorizations based on the work they will have to carry out [127].
- Communication. The organization communicates who are the authorized employees and monitors the various accesses [123]–[126].
- Distraction. When it comes to identity management and access control, possible mistakes caused by employee distraction should be considered. For example, since smart working is an increasingly adopted mode of work, organizations should have precise data access and management rules for employees working from home. In fact, the risk of a remote worker accidentally introducing malware connections into the company's computer network has been confirmed to be higher [159].

Awareness and Training (PR.AT):

"The organization's staff and partners receive cybersecurity awareness training and are trained to carry out their cybersecurity duties and responsibilities in accordance with relevant policies, procedures and agreements."

Human Factors:

- Awareness. Awareness campaigns can be very useful within an organization. By making everyone in the organization aware of the risks, they contribute to effective cybersecurity. It is recommended that organizations promote these campaigns clearly and engagingly [189], [190].
- Knowledge. to expand knowledge and ensure more effective cybersecurity, the need for training is emphasized. The company encourages training to expand knowledge and gain cybersecurity benefits. In this way, the employee can avoid cyber errors and breaches [147], [148]

- Norms. it's very important to educate the employee to broaden their knowledge and improve cybersecurity effectiveness. Making employees comfortable with cybersecurity policies and norms is part of employee training [149].
- Complacency. if the employee is completely disinterested in how he/she should act correctly in the company and has adverse behavior toward company training, it can be a sign of some desensitization to online risks. This situation can be attributed to employee complacency [107], [128], [133].
- Fatigue. If corporate training involves a high workload, the employee may feel fatigued and disinterested. In addition, because of the mandatory training, the employee may feel overwhelmed and deprived of his or her freedom generating unproductive behavior [128], [166].

Data Security (PR.DS):
"Information and records (data) are managed consistent with the organization's risk strategy to protect the confidentiality, integrity, and availability of information."

Human Factors:

- Norms. The challenge of designing security that is effective but usable is a core aspect of computer and information security. Research has demonstrated that users actively avoid security mechanisms that are difficult to use, and/or make mistakes that might undermine security [198]. Security must be user-centered, but implementing user-experience principles to improve usability is still an open issue regarding the current implementation of cybersecurity in organizations [108], [198], [199].
- Stress. If the procedures regarding data security are too difficult to manage the employee could experience stress. There is a need to reduce the complexity associated with technical terms and a large amount of information to be memorized [174].
- Complacency. The most confident workers from a cyber risk perspective are often those who occasionally act to innocently circumvent data security policies in the name of expediency or convenience. This is the case, for example, with users who violate security policies by sharing passwords, and not following security procedures to access email or hardware systems. The reasons that lead them to such behavior are often related to a desire to go faster, at the expense of security [198], [200], [201].
- Fatigue. An overwhelming experience induced by the practice of committing too many passwords to memory could cause security fatigue [183]. Cybersecurity professionals need to create password policies that are user-friendly and less frustrating. The disconnect between cybersecurity

professionals and users emphasizes the need for psychology-based professionals to develop less fatigue-inducing policies and practices [197].

Information Protection Processes and Procedures (PR.IP):

"Security policies (that address purpose, scope, roles, responsibilities, management commitment, and coordination among organizational entities), processes, and procedures are maintained and used to manage protection of information systems and assets."

Human Factors:

- Fatigue. If roles, responsibilities, and processes to follow are not clear employees may experience cybersecurity fatigue [128], [183].
- Norms. In the context of standards, it is important to clearly define the procedures to be followed and the roles to be respected. Everything must be clear, and user-friendly. The challenge is to design security norms that are effective and usable [108], [198], [199].
- Pressure. If the processes give employees a high workload or if the work demands from the organization do not meet the employee's abilities, they may suffer work pressure [154].
- Complacency. As stated in the previous point, complacent workers are more likely to find ways to speed up security procedures thus increasing system vulnerabilities [198], [200], [201].
- Stress. procedures regarding information protection should be accurately explained. When those results are too difficult to manage the employee could experience stress and increase the cyber risk [174].
- Communication. It is recommended to adopt Secure Communication Principles. NCSC guides with 7 principles, for risk owners and security professionals who wish to assess communication technologies for use in their organizations, to help them achieve the right balance of functionality, security and privacy. It is of relevance for those working in the public sector. Principles are: Protect data in transit; Protect network nodes with access to sensitive data; Protect against unauthorized user access to the service; Provision for secure audit of the service; Use metadata only for its necessary purpose; Assess supply chain for trust and resilience [202].

Maintenance (PR.MA):

"Maintenance and repairs of industrial control and information system components are performed consistent with policies and procedures. "

Human Factors:

- Resources. Maintenance is the core function to keep a system running and avoid failure. If the maintenance is not applied to all computer systems and resources the information system is more vulnerable and exposed to cyber attacks, system failures, slow running, and low quality [168].
- Stress. When an employee finds himself/herself working with outdated and damaged components, he/she perceives an increased possibility of error and risk, which generates a burden of stress in him/her [182], [183].

Protective Technology (PR.PT):

"Technical security solutions are managed to ensure the security and resilience of systems and assets, consistent with related policies, procedures, and agreements."

Human Factors:

- Resources. Special tamper-evident features and materials must be acquired by organizations in order to detect, tampering, and prevent efforts to compromise, change, penetrate, extract, or replace information processing equipment and keying material. To effectively safeguard information, it may be necessary to purchase a specific technology, as well as extra hardware, software, or services. Firewalls and intrusion detection systems are clear examples of items that would fit under this category.
- Norms. Norms must include the reviewing of audit/log records regularly to know who is accessing what and when. It is required to set the privileges based on need and monitor who has access to what and why they have that access [60].

Table 16 - NIST function PROTECT & Human Factors

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)	**Identity Management, Authentication and Access Control (PR.AC):** Access to physical and logical assets and associated facilities is limited to authorized users, processes, and devices, and is managed consistent with the assessed risk of unauthorized access to authorized activities and transactions.	**Fatigue**: Authentication fatigue is a condition of tiredness and fatigue due to the continuous validation of one's identity, accompanied by the requirement to create complicated passwords, and additional credentials to gain access **Norms:** The organization must give to its employees the authorizations based on the work they have to carry out **Communication**: The organization communicates who are the authorized employees and monitors the various accesses **Distraction**: When it comes to identity management and access control, possible mistakes of employee distraction must be considered.	**PR.AC-1:** Identities and credentials are issued, managed, verified, revoked, and audited for authorized devices, users and processes	**CIS CSC** 1, 5, 15, 16 **COBIT 5** DSS05.04, DSS06.03 **ISA 62443-2-1:2009** 4.3.3.5.1 **ISA 62443-3-3:2013** SR 1.1, SR 1.2, SR 1.3, SR 1.4, SR 1.5, SR 1.7, SR 1.8, SR 1.9 **ISO/IEC 27001:2013** A.9.2.1, A.9.2.2, A.9.2.3, A.9.2.4, A.9.2.6, A.9.3.1, A.9.4.2, A.9.4.3 **NIST SP 800-53 Rev. 4** AC-1, AC-2, IA-1, IA-2, IA-3, IA-4, IA-5, IA-6, IA-7, IA-8, IA-9, IA-10, IA-11
			PR.AC-2: Physical access to assets is managed and protected	**COBIT 5** DSS01.04, DSS05.05 **ISA 62443-2-1:2009** 4.3.3.3.2, 4.3.3.3.8 **ISO/IEC 27001:2013** A.11.1.1, A.11.1.2, A.11.1.3, A.11.1.4, A.11.1.5, A.11.1.6, A.11.2.1, A.11.2.3, A.11.2.5, A.11.2.6, A.11.2.7, A.11.2.8 **NIST SP 800-53 Rev. 4** PE-2, PE-3, PE-4, PE-5, PE-6, PE-8

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)			**PR.AC-3:** Remote access is managed	**CIS CSC** 12 **COBIT 5** APO13.01, DSS01.04, DSS05.03 **ISA 62443-2-1:2009** 4.3.3.6.6 **ISA 62443-3-3:2013** SR 1.13, SR 2.6 **ISO/IEC 27001:2013** A.6.2.1, A.6.2.2, A.11.2.6, A.13.1.1, A.13.2.1 **NIST SP 800-53 Rev. 4** AC-1, AC-17, AC-19, AC-20, SC-15
			PR.AC-4: Access permissions and authorizations are managed, incorporating the principles of least privilege and separation of duties	**CIS CSC** 3, 5, 12, 14, 15, 16, 18 **COBIT 5** DSS05.04 **ISA 62443-2-1:2009** 4.3.3.7.3 **ISA 62443-3-3:2013** SR 2.1 **ISO/IEC 27001:2013** A.6.1.2, A.9.1.2, A.9.2.3, A.9.4.1, A.9.4.4, A.9.4.5 **NIST SP 800-53 Rev. 4** AC-1, AC-2, AC-3, AC- 5, AC-6, AC-14, AC-16, AC-24
			PR.AC-5: Network integrity is protected (e.g., network segregation, network segmentation)	**CIS CSC** 9, 14, 15, 18 **COBIT 5** DSS01.05, DSS05.02 **ISA 62443-2-1:2009** 4.3.3.4 **ISA 62443-3-3:2013** SR 3.1, SR 3.8 **ISO/IEC 27001:2013** A.13.1.1, A.13.1.3, A.13.2.1, A.14.1.2, A.14.1.3 **NIST SP 800-53 Rev. 4** AC-4, AC-10, SC-7
			PR.AC-6: Identities are proofed and bound to credentials and asserted in interactions	**CIS CSC**, 16 **COBIT 5** DSS05.04, DSS05.05, DSS05.07, DSS06.03 **ISA 62443-2-1:2009** 4.3.3.2.2, 4.3.3.5.2, 4.3.3.7.2, 4.3.3.7.4 **ISA 62443-3-3:2013** SR 1.1, SR 1.2, SR 1.4, SR 1.5, SR 1.9, SR 2.1 **ISO/IEC 27001:2013**, A.7.1.1, A.9.2.1 **NIST SP 800-53 Rev. 4** AC-1, AC-2, AC-3, AC- 16, AC-19, AC-24, IA-1, IA-2, IA-4, IA-5, IA-8, PE-2, PS-3

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)			**PR.AC-7:** Users, devices, and other assets are authenticated (e.g., single-factor, multi-factor) commensurate with the risk of the transaction (e.g., individuals' security and privacy risks and other organizational risks)	**CIS CSC** 1, 12, 15, 16 **COBIT 5** DSS05.04, DSS05.10, DSS06.10 **ISA 62443-2-1:2009** 4.3.3.6.1, 4.3.3.6.2, 4.3.3.6.3, 4.3.3.6.4, 4.3.3.6.5, 4.3.3.6.6, 4.3.3.6.7, 4.3.3.6.8, 4.3.3.6.9 **ISA 62443-3-3:2013** SR 1.1, SR 1.2, SR 1.5, SR 1.7, SR 1.8, SR 1.9, SR 1.10 **ISO/IEC 27001:2013** A.9.2.1, A.9.2.4, A.9.3.1, A.9.4.2, A.9.4.3, A.18.1.4 **NIST SP 800-53 Rev. 4** AC-7, AC-8, AC-9, AC- 11, AC-12, AC-14, IA-1, IA-2, IA-3, IA-4, IA-5, IA-8, IA-9, IA-10, IA-11
	Awareness and Training (PR.AT): The organization's personnel and partners are provided cybersecurity awareness education and are trained to perform their cybersecurity-related duties and responsibilities consistent with related policies, procedures, and agreements.	**Awareness:** Awareness campaigns can be very useful within an organization. By making everyone in the organization aware of the risks, they contribute to effective cybersecurity **Knowledge:** The company encourages training to expand knowledge and gain cybersecurity benefits **Norms:** Making employees comfortable with cybersecurity policies is part of employee training	**PR.AT-1:** All users are informed and trained	**CIS CSC** 17, 18 **COBIT 5** APO07.03, BAI05.07 **ISA 62443-2-1:2009** 4.3.2.4.2 **ISO/IEC 27001:2013** A.7.2.2, A.12.2.1 **NIST SP 800-53 Rev. 4** AT-2, PM-13

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)		**Fatigue:** If corporate training involves a high workload, the employee may feel fatigued and disinterested	**PR.AT-2:** Privileged users understand their roles and responsibilities	**CIS CSC** 5, 17, 18 **COBIT 5** APO07.02, DSS05.04, DSS06.03 **ISA 62443-2-1:2009** 4.3.2.4.2, 4.3.2.4.3 **ISO/IEC 27001:2013** A.6.1.1, A.7.2.2 **NIST SP 800-53 Rev. 4** AT-3, PM-13
		Complacency: If the employee is completely disinterested in how he/she should act correctly in the company and has adverse behavior toward company training, it can be a sign of some desensitization to online risks. This situation can be attributed to employee complacency	**PR.AT-3:** Third-party stakeholders (e.g., suppliers, customers, partners) understand their roles and responsibilities	**CIS CSC** 17 **COBIT 5** APO07.03, APO07.06, APO10.04, APO10.05 **ISA 62443-2-1:2009** 4.3.2.4.2 **ISO/IEC 27001:2013** A.6.1.1, A.7.2.1, A.7.2.2 **NIST SP 800-53 Rev. 4** PS-7, SA-9, SA-16
			PR.AT-4: Senior executives understand their roles and responsibilities	**CIS CSC** 17, 19 **COBIT 5** EDM01.01, APO01.02, APO07.03 **ISA 62443-2-1:2009** 4.3.2.4.2 **ISO/IEC 27001:2013** A.6.1.1, A.7.2.2 **NIST SP 800-53 Rev. 4** AT-3, PM-13
			PR.AT-5: Physical and cybersecurity personnel understand their roles and responsibilities	**CIS CSC** 17 **COBIT 5** APO07.03 **ISA 62443-2-1:2009** 4.3.2.4.2 **ISO/IEC 27001:2013** A.6.1.1, A.7.2.2
	Data Security (PR.DS): Information and records (data) are managed consistent with the organization's risk strategy to protect the confidentiality, integrity, and	**Norms:** The challenge of designing security that is effective but usable is a core aspect of the computer and information security **Stress:** If the procedures regarding data security are too	**PR.DS-1:** Data-at-rest is protected	**CIS CSC** 13, 14 **COBIT 5** APO01.06, BAI02.01, BAI06.01, DSS04.07, DSS05.03, DSS06.06 **ISA 62443-3-3:2013** SR 3.4, SR 4.1 **ISO/IEC 27001:2013** A.8.2.3 **NIST SP 800-53 Rev. 4** MP-8, SC-12, SC-28

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)	availability of information	difficult to manage the employee could experience stress. **Fatigue:** An overwhelming experience induced by the practice of committing too many passwords to memory could case security-fatigue **Complacency:** The most cyber risk confident workers are often the ones who occasionally act to innocently circumvent data security policies in the name of expediency or convenience.	**PR.DS-2:** Data-in-transit is protected	**CIS CSC** 13, 14 **COBIT 5** APO01.06, DSS05.02, DSS06.06 **ISA 62443-3-3:2013** SR 3.1, SR 3.8, SR 4.1, SR 4.2 **ISO/IEC 27001:2013** A.8.2.3, A.13.1.1, A.13.2.1, A.13.2.3, A.14.1.2, A.14.1.3 **NIST SP 800-53 Rev. 4** SC-8, SC-11, SC-12
			PR.DS-3: Assets are formally managed throughout removal, transfers, and disposition	**CIS CSC** 1 **COBIT 5** BAI09.03 **ISA 62443-2-1:2009** 4.3.3.3.9, 4.3.4.4.1 **ISA 62443-3-3:2013** SR 4.2 **ISO/IEC 27001:2013** A.8.2.3, A.8.3.1, A.8.3.2, A.8.3.3, A.11.2.5, A.11.2.7 **NIST SP 800-53 Rev. 4** CM-8, MP-6, PE-16
			PR.DS-4: Adequate capacity to ensure availability is maintained	**CIS CSC** 1, 2, 13 **COBIT 5** APO13.01, BAI04.04 **ISA 62443-3-3:2013** SR 7.1, SR 7.2 **ISO/IEC 27001:2013** A.12.1.3, A.17.2.1 **NIST SP 800-53 Rev. 4** AU-4, CP-2, SC-5
			PR.DS-5: Protections against data leaks are implemented	**CIS CSC** 13 **COBIT 5** APO01.06, DSS05.04, DSS05.07, DSS06.02 **ISA 62443-3-3:2013** SR 5.2 **ISO/IEC 27001:2013** A.6.1.2, A.7.1.1, A.7.1.2, A.7.3.1, A.8.2.2, A.8.2.3, A.9.1.1, A.9.1.2, A.9.2.3, A.9.4.1, A.9.4.4, A.9.4.5, A.10.1.1, A.11.1.4, A.11.1.5, A.11.2.1, A.13.1.1, A.13.1.3, A.13.2.1, A.13.2.3, A.13.2.4, A.14.1.2, A.14.1.3 NIST SP 800-53 Rev. 4 AC-4, AC-5, AC-6, PE- 19, PS-3, PS-6, SC-7, SC-8, SC-13, SC-31, SI-4

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)	availability of information		**PR.DS-6:** Integrity checking mechanisms are used to verify software, firmware, and information integrity	**CIS CSC** 2, 3 **COBIT 5** APO01.06, BAI06.01, DSS06.02 **ISA 62443-3-3:2013** SR 3.1, SR 3.3, SR 3.4, SR 3.8 **ISO/IEC 27001:2013** A.12.2.1, A.12.5.1, A.14.1.2, A.14.1.3, A.14.2.4 **NIST SP 800-53 Rev. 4** SC-16, SI-7
			PR.DS-7: The development and testing environment(s) are separate from the production environment	**CIS CSC** 18, 20 **COBIT 5** BAI03.08, BAI07.04 **ISO/IEC 27001:2013** A.12.1.4 **NIST SP 800-53 Rev. 4** CM-2
			PR.DS-8: Integrity checking mechanisms are used to verify hardware integrity	**COBIT 5** BAI03.05 **ISA 62443-2-1:2009** 4.3.4.4.4 **ISO/IEC 27001:2013** A.11.2.4 **NIST SP 800-53 Rev. 4** SA-10, SI-7
	Information Protection Processes and Procedures (PR.IP): Security policies (that address purpose, scope, roles, responsibilities, management commitment, and coordination among organizational entities), processes, and procedures are maintained	**Fatigue:** If roles and responsibilities are not clear employees may experience cybersecurity fatigue **Norms:** In the context of standards, it is important to clearly define the procedures to be followed and the roles to be respected. Everything must be clear, and user-friendly.	**PR.IP-1:** A baseline configuration of information technology/industrial control systems is created and maintained incorporating security principles (e.g. concept of least functionality)	**CIS CSC** 3, 9, 11 **COBIT 5** BAI10.01, BAI10.02, BAI10.03, BAI10.05 **ISA 62443-2-1:2009** 4.3.4.3.2, 4.3.4.3.3 **ISA 62443-3-3:2013** SR 7.6 **ISO/IEC 27001:2013** A.12.1.2, A.12.5.1, A.12.6.2, A.14.2.2, A.14.2.3, A.14.2.4 **NIST SP 800-53 Rev. 4** CM-2, CM-3, CM-4, CM- 5, CM-6, CM-7, CM-9, SA-10

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)	and used to manage protection of information systems and assets.	**Pressure**: If the processes give employees a high workload or if the work demands from the organization do not meet the employee's abilities, they may suffer work pressure **Complacency:** complacent workers are more likely to find ways to speed up security procedures thus increasing system vulnerabilities **Stress:** procedures regarding information protection should be accurately explained **Communication:** Adopt secure communication principles	**PR.IP-2:** A System Development Life Cycle to manage systems is implemented	**CIS CSC** 18 **COBIT 5** APO13.01, BAI03.01, BAI03.02, BAI03.03 **ISA 62443-2-1:2009** 4.3.4.3.3 **ISO/IEC 27001:2013** A.6.1.5, A.14.1.1, A.14.2.1, A.14.2.5 **NIST SP 800-53 Rev. 4** PL-8, SA-3, SA-4, SA-8, SA-10, SA-11, SA-12, SA-15, SA-17, SI-12, SI- 13, SI-14, SI-16, SI-17
			PR.IP-3: Configuration change control processes are in place	**CIS CSC** 3, 11 **COBIT 5** BAI01.06, BAI06.01 **ISA 62443-2-1:2009** 4.3.4.3.2, 4.3.4.3.3 **ISA 62443-3-3:2013** SR 7.6 **ISO/IEC 27001:2013** A.12.1.2, A.12.5.1, A.12.6.2, A.14.2.2, A.14.2.3, A.14.2.4 **NIST SP 800-53 Rev. 4** CM-3, CM-4, SA-10
			PR.IP-4: Backups of information are conducted, maintained, and tested	**CIS CSC** 10 **COBIT 5** APO13.01, DSS01.01, DSS04.07 **ISA 62443-2-1:2009** 4.3.4.3.9 **ISA 62443-3-3:2013** SR 7.3, SR 7.4 **ISO/IEC 27001:2013** A.12.3.1, A.17.1.2, A.17.1.3, A.18.1.3 **NIST SP 800-53 Rev. 4** CP-4, CP-6, CP-9
			PR.IP-5: Policy and regulations regarding the physical operating environment for organizational assets are met	**COBIT 5** DSS01.04, DSS05.05 **ISA 62443-2-1:2009** 4.3.3.3.1 4.3.3.3.2, 4.3.3.3.3, 4.3.3.3.5, 4.3.3.3.6 **ISO/IEC 27001:2013** .11.1.4, A.11.2.1, A.11.2.2, A.11.2.3 **NIST SP 800-53 Rev. 4** PE-10, PE-12, PE-13, PE- 14, PE-15, PE-18

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)			**PR.IP-6:** Data is destroyed according to policy	**COBIT 5** BAI09.03, DSS05.06 **ISA 62443-2-1:2009** 4.3.4.4.4 **ISA 62443-3-3:2013** SR 4.2 **ISO/IEC 27001:2013** A.8.2.3, A.8.3.1, A.8.3.2, A.11.2.7 **NIST SP 800-53 Rev. 4** MP-6
			PR.IP-7: Protection processes are improved	**COBIT 5** APO11.06, PO12.06, DSS04.05 **ISA 62443-2-1:2009** 4.4.3.1, 4.4.3.2, 4.4.3.3, 4.4.3.4, 4.4.3.5, 4.4.3.6, 4.4.3.7, 4.4.3.8 **ISO/IEC 27001:2013** A.16.1.6, Clause 9, Clause 10 **NIST SP 800-53 Rev. 4** CA-2, CA-7, CP-2, IR-8, PL-2, PM-6
			PR.IP-8: Effectiveness of protection technologies is shared	**COBIT 5** BAI08.04, DSS03.04 **ISO/IEC 27001:2013** A.16.1.6 **NIST SP 800-53 Rev. 4** AC-21, CA-7, SI-4
			PR.IP-9: Response plans (Incident Response and Business Continuity) and recovery plans (Incident Recovery and Disaster Recovery) are in place and managed	**CIS CSC** 19 **COBIT 5** APO12.06, DSS04.03 **ISA 62443-2-1:2009** 4.3.2.5.3, 4.3.4.5.1 **ISO/IEC 7001:2013** A.16.1.1, A.17.1.1, A.17.1.2, A.17.1.3 **NIST SP 800-53 Rev. 4** CP-2, CP-7, CP-12, CP- 13, IR-7, IR-8, IR-9, PE-17
			PR.IP-10: Response and recovery plans are tested	**CIS CSC** 19, 20 **COBIT 5** DSS04.04 **ISA 62443-2-1:2009** 4.3.2.5.7, .3.4.5.11 **ISA 62443-3-3:2013** SR 3.3 **ISO/IEC 27001:2013** A.17.1.3 **NIST SP 800-53 Rev. 4** CP-4, IR-3, PM-14

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)			**PR.IP-11:** Cybersecurity is included in human resources practices (e.g., deprovisioning, personnel screening)	**CIS CSC** 5, 16 **COBIT 5** APO07.01, APO07.02, APO07.03, APO07.04, APO07.05 **ISA 62443-2-1:2009** 4.3.3.2.1, 4.3.3.2.2, 4.3.3.2.3 **ISO/IEC 7001:2013** A.7.1.1, A.7.1.2, A.7.2.1, A.7.2.2, A.7.2.3, A.7.3.1, A.8.1.4 **NIST SP 800-53 Rev. 4** PS-1, PS-2, PS-3, PS-4, PS-5, PS-6, PS-7, PS-8, SA-21
			PR.IP-12: A vulnerability management plan is developed and implemented	**CIS CSC** 4, 18, 20 **COBIT 5** BAI03.10, DSS05.01, DSS05.02 **ISO/IEC 27001:2013** A.12.6.1, A.14.2.3, A.16.1.3, A.18.2.2, A.18.2.3 **NIST SP 800-53 Rev. 4** RA-3, RA-5, SI-2
	Maintenance (PR.MA): Maintenance and repairs of industrial control and information system components are performed consistent with policies and procedures.	**Resources:** If the maintenance is not applied to all computer systems and resources the information system is more vulnerable and exposed to cyber attacks, system failures, slow running, and low quality **Stress:** When an employee finds himself/herself working with outdated and damaged components, he/she perceives an increased possibility of error and risk, which generates a burden of stress in him/her	**PR.MA-1:** Maintenance and repair of organizational assets are performed and logged, with approved and controlled tools	**COBIT 5** BAI03.10, BAI09.02, BAI09.03, DSS01.05 **ISA 62443-2-1:2009** 4.3.3.3.7 **ISO/IEC 27001:2013** A.11.1.2, A.11.2.4, A.11.2.5, A.11.2.6 **NIST SP 800-53 Rev. 4** MA-2, MA-3, MA-5, MA-6

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)			**PR.MA-2:** Remote maintenance of organizational assets is approved, logged, and performed in a manner that prevents unauthorized access	**CIS CSC** 3, 5 **COBIT 5** DSS05.04 **ISA 62443-2-1:2009** 4.3.3.6.5, 4.3.3.6.6, 4.3.3.6.7, 4.3.3.6.8 **ISO/IEC 27001:2013** A.11.2.4, A.15.1.1, A.15.2.1 **NIST SP 800-53 Rev. 4** MA-4
	Protective Technology (PR.PT): Technical security solutions are managed to ensure the security and resilience of systems and assets, consistent with related policies, procedures, and agreements.	**Resources:** Special tamper-evident features and materials must be acquired by organizations in order to detect, tampering, and prevent efforts to compromise, change, penetrate, extract, or replace information processing equipment and keying material. **Norms:** Norms must include the reviewing of audit/log records regularly to know who is accessing what and when. It is required to set the privileges based on need and monitor who has access to what and why they have that access	**PR.PT-1:** Audit/log records are determined, documented, implemented, and reviewed in accordance with policy	**CIS CSC** 1, 3, 5, 6, 14, 15, 16 **COBIT 5** APO11.04, BAI03.05, DSS05.04, DSS05.07, MEA02.01 **ISA 62443-2-1:2009** 4.3.3.3.9, 4.3.3.5.8, 4.3.4.4.7, 4.4.2.1, 4.4.2.2, 4.4.2.4 **ISA 62443-3-3:2013** SR 2.8, SR 2.9, SR 2.10, SR 2.11, SR 2.12 **ISO/IEC 27001:2013** A.12.4.1, A.12.4.2, A.12.4.3, A.12.4.4, A.12.7.1 **NIST SP 800-53 Rev. 4** AU Family
			PR.PT-2: Removable media is protected and its use restricted according to policy	**CIS CSC** 8, 13 **COBIT 5** APO13.01, DSS05.02, DSS05.06 **ISA 62443-3-3:2013** SR 2.3 **ISO/IEC 27001:2013** A.8.2.1, A.8.2.2, A.8.2.3, A.8.3.1, A.8.3.3, A.11.2.9 **NIST SP 800-53 Rev. 4** MP-2, MP-3, MP-4, MP- 5, MP-7, MP-8

Function	Category	Human Factors	Subcategory	Informative References
PROTECT (PR)			**PR.PT-3:** The principle of least functionality is incorporated by configuring systems to provide only essential capabilities	**CIS CSC** 3, 11, 14 **COBIT 5** DSS05.02, SS05.05, SS06.06 **ISA 62443-2-1:2009** 4.3.3.5.1, 4.3.3.5.2, 4.3.3.5.3, 4.3.3.5.4, 4.3.3.5.5, 4.3.3.5.6, 4.3.3.5.7, 4.3.3.5.8, 4.3.3.6.1, 4.3.3.6.2, 4.3.3.6.3, 4.3.3.6.4, 4.3.3.6.5, 4.3.3.6.6, 4.3.3.6.7, 4.3.3.6.8, 4.3.3.6.9, 4.3.3.7.1, 4.3.3.7.2, 4.3.3.7.3, 4.3.3.7.4 **ISA 2443-3-3:2013** SR 1.1, SR 1.2, SR 1.3, SR 1.4, SR 1.5, SR 1.6, SR 1.7, SR 1.8, SR 1.9, SR 1.10, SR 1.11, SR 1.12, SR 1.13, SR 2.1, SR 2.2,SR 2.3, SR 2.4, SR 2.5, SR 2.6, SR 2.7 **ISO/IEC 27001:2013** A.9.1.2 **NIST SP 800-53 Rev. 4** AC-3, CM-7
			PR.PT-4: Communications and control networks are protected	**CIS CSC** 8, 12, 15 **COBIT 5** DSS05.02, APO13.01 **ISA 62443-3-3:2013** SR 3.1, SR 3.5, SR 3.8, SR 4.1, SR 4.3, SR 5.1, SR 5.2, SR 5.3, SR 7.1, SR 7.6 **ISO/IEC 7001:2013** A.13.1.1, A.13.2.1, A.14.1.3 **NIST SP 800-53 Rev. 4** AC-4, AC-17, AC-18, CP-8, SC-7, SC-19, SC-20, SC-21, SC-22, SC-23, SC-24, SC-25, SC-29, SC-32, SC-36, SC-37, SC- 38, SC-39, SC-40, SC-41, SC-43
			PR.PT-5: Mechanisms (e.g., failsafe, load balancing, hot swap) are implemented to achieve resilience requirements in normal and adverse situations	**COBIT 5** BAI04.01, BAI04.02, BAI04.03, BAI04.04, BAI04.05, DSS01.05 **ISA 62443-2-1:2009** 4.3.2.5.2 **ISA 62443-3-3:2013** SR 7.1, SR 7.2 **ISO/IEC 27001:2013** A.17.1.2, A.17.2.1 **NIST SP 800-53 Rev. 4** CP-7, CP-8, CP-11, CP- 13, PL-8, SA-14, SC-6

3. DETECT

Develop and implement appropriate activities to identify the occurrence of a cybersecurity event.

Anomalies and Events (DE.AE):

"Anomalous activity is detected and the potential impact of event is understood."

Human Factors:

- Awareness. Courses on awareness are meant to assist people in spotting cyberattacks and taking the necessary countermeasures. for this reason, a company needs to motivate its employees and praise them when they detect and mitigate an attack [140], [141], [154]
- Knowledge. Cyber-hygiene promotes expert consensus on important harmful threats to the organization, crucial safeguarding behaviors to protect environments, and risky crucial user behavior to support cybersecurity environments. In order to comprehend and use appropriate cyber-hygiene practices, training is required [147], [150]
- Resources. By adopting a cybersecurity framework, the organization can develop guidelines to be better equipped for anomalies activities detection [167], [203].
- Distraction. employees can get distracted during the anomaly detection phase as they believe that the system can automatically detect the threat. This overtrust in cybersecurity devices can mean that anomalies not detected by the system can enter and create serious problems [94].
- Pressure. People who work in this area are faced with overwhelming pressure. Errors in their judgment, caused by excessive work-related pressure can indeed have detrimental effects on business and personal data [173].
- Teamwork. especially when we are talking about detection, team interaction could lead to better cyber defense performance [204].

Security Continuous Monitoring (DE.CM):

"The information system and assets are monitored to identify cybersecurity events and verify the effectiveness of protective measures"

Human Factors:

- Communication. To verify the effectiveness of the measures of the organization, it can be useful to apply a good feedback exchange system. Communication is a crucial component of developing a collaborative culture, and through the continuous use of feedback, by improving the personal performance of the employees, the overall performance of the organization is improved [127].

- Complacency. When human operators complacently oversee automation, they are less prepared to manually take control when their intervention is needed. Complacency may be related to the development of inappropriate trust [137].
- Distraction. Employees during the monitoring can overtrust cybersecurity devices which can lead to poor control and mismanagement of security measures [136].
- Fatigue. Especially in the monitoring phase, companies are increasingly relying on threat detection software. This software generates alerts and it is then up to the cyber analysts to verify whether the alert produced is a real threat. Unfortunately, many of these alerts are unfounded and cause so-called threat alert fatigue. In practice, there are more alerts than cyber analysts can properly investigate. This leads to an information overload problem where cyber analysts miss true attack alerts in the noise of false alarms [205].

Detection Processes (DE.DP):

"Detection processes and procedures are maintained and tested to ensure awareness of anomalous events."

Human Factors:

- Awareness. To be detected, the employee must be aware of the risk. For this reason, a company needs to organize Awareness Campaigns but it's not enough. To encourage active participation in these campaigns, the organization must motivate its employees and praise them when they detect and mitigate an attack [60], [140], [141], [154], [189].
- Knowledge. Training and cyber-hygiene can help employees detect anomalies. Cyber-hygiene promotes expert consensus on important harmful threats to the organization, crucial safeguarding behaviors to protect environments, and risky crucial user behavior to support cybersecurity environments. In order to comprehend and use appropriate cyber-hygiene practices, training is required [147], [148], [150].
- Resources. By adopting a cybersecurity framework, the organization can develop guidelines to better build and test detection processes and procedures [167], [203].
- Distraction. Employees can get distracted during the anomaly detection phase as they believe that the system can automatically detect the threat. This overtrust in cybersecurity devices can mean that anomalies not detected by the system can enter and create serious problems [136].

- Pressure. People who work in this area are faced with overwhelming pressure. Errors in their judgment, caused by excessive work-related pressure can indeed have detrimental effects on business and personal data [173].
- Teamwork: Especially when we are talking about detection, team interaction could lead to better cyber defense performance [204].

Table 17 - NIST function DETECT & Human Factors

Function	Category	Human Factors	Subcategory	Informative References
DETECT (DE)	**Anomalies and Events (DE.AE):** Anomalous activity is detected and the potential impact of event is understood	**Awareness:** Courses on awareness can assist people in spotting cyberattacks and taking measures **Knowledge:** Training and cyber-hygiene can help employees detect anomalies **Distraction:** employees can get distracted during the detection phase as they believe that the system can automatically detect the threat **Pressure:** People who work in detection are faced with overwhelming pressure. Errors in their judgment, can have detrimental effects on business and personal data **Teamwork:** team interaction could lead to better cyber defense performance	**DE.AE-1:** A baseline of network operations and expected data flows for users and systems is established and managed	**CIS CSC** 1, 4, 6, 12, 13, 15, 16 **COBIT 5** DSS03.01 **ISA 62443-2-1:2009** 4.4.3.3 **ISO/IEC 27001:2013** A.12.1.1, A.12.1.2, A.13.1.1, A.13.1.2 **NIST SP 800-53 Rev. 4** AC-4, CA-3, CM-2, SI-4

Function	Category	Human Factors	Subcategory	Informative References
DETECT (DE)			**DE.AE-2:** Detected events are analyzed to understand attack targets and methods	**CIS CSC** 3, 6, 13, 15 **COBIT 5** DSS05.07 **ISA 62443-2-1:2009** 4.3.4.5.6, 4.3.4.5.7, 4.3.4.5.8 **ISA 62443-3-3:2013** SR 2.8, SR 2.9, SR 2.10, SR 2.11, SR 2.12, SR 3.9, SR 6.1, SR 6.2 **ISO/IEC 27001:2013** A.12.4.1, A.16.1.1, A.16.1.4 **NIST SP 800-53 Rev. 4** AU-6, CA-7, IR-4, SI-4
			DE.AE-3: Event data are collected and correlated from multiple sources and sensors	**CIS CSC** 1, 3, 4, 5, 6, 7, 8, 11, 12, 13, 14, 15, 16 **COBIT 5** BAI08.02 **ISA 62443-3-3:2013** SR 6.1 **ISO/IEC 27001:2013** A.12.4.1, A.16.1.7 **NIST SP 800-53 Rev. 4** AU-6, CA-7, IR-4, IR-5, IR-8, SI-4
			DE.AE-4: Impact of events is determined	**CIS CSC** 4, 6 **COBIT 5** APO12.06, DSS03.01 **ISO/IEC 27001:2013** A.16.1.4 **NIST SP 800-53 Rev. 4** CP-2, IR-4, RA-3, SI-4
			DE.AE-5: Incident alert thresholds are established	**CIS CSC** 6, 19 **COBIT 5** APO12.06, DSS03.01 **ISA 62443-2-1:2009** 4.2.3.10 **ISO/IEC 27001:2013** A.16.1.4 **NIST SP 800-53 Rev. 4** IR-4, IR-5, IR-8
	Security Continuous Monitoring (DE.CM): The information system and assets are monitored to identify cybersecurity events and	**Communication:** to verify the effectiveness of the measures of the organization, it can be useful to apply a good feedback exchange system	**DE.CM-1:** The network is monitored to detect potential cybersecurity events	**CIS CSC** 1, 7, 8, 12, 13, 15, 16 **COBIT 5** DSS01.03, DSS03.05, DSS05.07 **ISA 62443-3-3:2013** SR 6.2 **NIST SP 800-53 Rev. 4** AC-2, AU-12, CA-7, CM- 3, SC-5, SC-7, SI-4

Function	Category	Human Factors	Subcategory	Informative References
DETECT (DE)	verify the effectiveness of protective measures	**Complacency:** When human operators complacently oversee automation, they are less prepared to manually take control when their intervention is needed. Complacency may be related to the development of inappropriate trust **Distraction:** employee during the monitoring can overtrust in cybersecurity devices which can lead to poor control and mismanagement of security measures **Fatigue:** cyber analysts should verify whether the alert produced is a real threat. Unfortunately, many of these alerts are unfounded and cause the so-called threat alert fatigue	**DE.CM-2:** The physical environment is monitored to detect potential cybersecurity events	**COBIT 5** DSS01.04, DSS01.05 **ISA 62443-2-1:2009** 4.3.3.3.8 **ISO/IEC 27001:2013** A.11.1.1, A.11.1.2 **NIST SP 800-53 Rev. 4** CA-7, PE-3, PE-6, PE-20
			DE.CM-3: Personnel activity is monitored to detect potential cybersecurity events	**CIS CSC** 5, 7, 14, 16 **COBIT 5** DSS05.07 **ISA 62443-3-3:2013** SR 6.2 **ISO/IEC 27001:2013** A.12.4.1, A.12.4.3 **NIST SP 800-53 Rev. 4** AC-2, AU-12, AU-13, CA-7, CM-10, CM-11
			DE.CM-4: Malicious code is detected	**CIS CSC** 4, 7, 8, 12 **COBIT 5** DSS05.01 **ISA 62443-2-1:2009** 4.3.4.3.8 **ISA 62443-3-3:2013** SR 3.2 **ISO/IEC 27001:2013** A.12.2.1 **NIST SP 800-53 Rev. 4** SI-3, SI-8
			DE.CM-5: Unauthorized mobile code is detected	**CIS CSC** 7, 8 **COBIT 5** DSS05.01 **ISA 62443-3-3:2013** SR 2.4 **ISO/IEC 27001:2013** A.12.5.1, A.12.6.2 **NIST SP 800-53 Rev. 4** SC-18, SI-4, SC-44

Function	Category	Human Factors	Subcategory	Informative References
DETECT (DE)			**DE.CM-6:** External service provider activity is monitored to detect potential cybersecurity events	**COBIT 5** APO07.06, APO10.05 **ISO/IEC 27001:2013** A.14.2.7, A.15.2.1 **NIST SP 800-53 Rev. 4** CA-7, PS-7, SA-4, SA-9, SI-4
			DE.CM-7: Monitoring for unauthorized personnel, connections, devices, and software is performed	**CIS CSC** 1, 2, 3, 5, 9, 12, 13, 15, 16 **COBIT 5** DSS05.02, DSS05.05 **ISO/IEC 27001:2013** A.12.4.1, A.14.2.7, A.15.2.1 **NIST SP 800-53 Rev. 4** AU-12, CA-7, CM-3, CM-8, PE-3, PE-6, PE-20, SI-4
			DE.CM-8: Vulnerability scans are performed	**CIS CSC** 4, 20 **COBIT** 5 BAI03.10, DSS05.01 **ISA 62443-2-1:2009** 4.2.3.1, 4.2.3.7 **ISO/IEC 27001:2013** A.12.6.1 **NIST** SP 800-53 Rev. 4 RA-5
	Detection Processes (DE.DP): Detection processes and procedures are maintained and tested to ensure awareness of anomalous events.	**Resources:** by adopting a framework, guidelines can be defined for detect anomalies and events **Distraction:** During the work the employee can have an overtrust in cybersecurity devices which can	**DE.DP-1:** Roles and responsibilities for detection are well defined to ensure accountability	**CIS CSC** 19 **COBIT 5** APO01.02**,** DSS05.01, DSS06.03 **ISA 62443-2-1:2009** 4.4.3.1 **ISO/IEC 27001:2013** A.6.1.1, A.7.2.2 **NIST SP 800-53 Rev. 4** CA-2, CA-7, PM-14
			DE.DP-2: Detection activities comply with all applicable requirements	**COBIT 5** DSS06.01, MEA03.03, MEA03.04 **ISA 62443-2-1:2009** 4.4.3.2 **ISO/IEC 27001:2013** A.18.1.4, A.18.2.2, A.18.2.3 **NIST SP 800-53 Rev. 4** AC-25, CA-2, CA-7, SA- 18, SI-4, PM-14

Function	Category	Human Factors	Subcategory	Informative References
DETECT (DE)		lead to problems if an attack is not detected automatically **Awareness:** To be detected, the employee must be aware of the risk **Knowledge:** training and cyber-hygiene can help employees detect anomalies **Teamwork:** especially when we are talking about detection, team interaction could lead to better cyber defense performance **Pressure:** People who work in this area are faced with overwhelming pressure. Errors in their judgment, caused by excessive work-related stress can indeed have detrimental effects upon business and personal data	**DE.DP-3:** Detection processes are tested	**COBIT 5** APO13.02, DSS05.02 **ISA 62443-2-1:2009** 4.4.3.2 **ISA 62443-3-3:2013** SR 3.3 **ISO/IEC 27001:2013** A.14.2.8 **NIST SP 800-53 Rev. 4** CA-2, CA-7, PE-3, SI-3, SI-4, PM-14
			DE.DP-4: Event detection information is communicated	**CIS CSC** 19 **COBIT 5** APO08.04, APO12.06, DSS02.05 **ISA 62443-2-1:2009** 4.3.4.5.9 **ISA 62443-3-3:2013** SR 6.1 **ISO/IEC 27001:2013** A.16.1.2, A.16.1.3 **NIST SP 800-53 Rev. 4** AU-6, CA-2, CA-7, RA- 5, SI-4
			DE.DP-5: Detection processes are continuously improved	**COBIT 5** APO11.06, APO12.06, DSS04.05 **ISA 62443-2-1:2009** 4.4.3.4 **ISO/IEC 27001:2013** A.16.1.6 **NIST SP 800-53 Rev. 4**, CA-2, CA-7, PL-2, RA- 5, SI-4, PM-14

4. RESPOND

Develop and implement appropriate activities to take action regarding a detected cybersecurity incident.

Response Planning (RS.RP):

"Response processes and procedures are executed and maintained, to ensure response to detected cybersecurity incidents."

Human Factors:

- Distraction. If the working environment is inadequate and the workload is extremely heavy, distraction-dependent errors may occur when conducting and sustaining response procedures [116], [159].
- Fatigue. Involuntary errors caused by fatigue may occur during response procedures and maintenance of ongoing processes if the employee has a heavy task and a long shift [132].
- Knowledge and Awareness. Stakeholders must be prepared for and respond to security threats by implementing stringent yet multi-faceted cybersecurity measures such as security awareness (SA) and incident response (IR) training. The difficulty is that most SA and IR training has been criticized for being overly conventional, with participants reading newsletters, posters, brochures, quizzes, and handouts with information on critical security matters and answering questions. Opponents of this sort of training contend that it is a tedious, time-consuming, and uninteresting tick-box activity. In response to these critiques, several stakeholders have frequently stated that IR training offered in this manner is costly and that it diverts participants' attention away from productive and business-sustaining tasks. One possible option for making training more enjoyable is to ensure that training is structured to be practical, using training activities such as role-playing, games, virtual reality, and simulation exercises [186], [206], [207].
- Pressure. Employees may experience Time Pressure because there is a time limit on the response tasks that must be carried out [154].
- Complacency. During the response process, people's seeming complacency is mostly due to the lack of a significant and damaging cyber incident that they had dealing with or because they shift away from self-responsibility during work, sure that "others" (whether other people or other technologies) have taken care of the user's system's security without the need for personal involvement [134], [138].

Communications (RS.CO):

"Response activities are coordinated with internal and external stakeholders (e.g. external support from law enforcement agencies)."

Human Factors:

- Teamwork. During team communication, the employees must show respect for the other people to improve the teamwork. The organization must improve a good direct and indirect communication channel to coordinate the work with the other stakeholders [162], [163].
- Communication. When in contact with several people, the required soft skills are very important so that interpersonal conflict does not sour. When implementing a cooperation plan, employees must have the required soft skills to avoid problems and slowdowns during communication [130], [189].
- Stress. If two or more interacting people can't get along, don't agree, or are otherwise different, there may be interpersonal conflicts that generate general stress [130], [154], [173]. There may be differences between the personal traits of the people who must collaborate.
- Fatigue. Communication with various entities can cause cognitive fatigue and tiredness. For example, an employee who is tired of being told what to do may feel exhausted and stressed due to the extreme pressure and over-supervision within the workplace [128].
- Assertiveness. In conversations between employees and employers, managers and supervisors, assertiveness is crucial. A deficiency in this skill can undermine employees' job performance. Assertiveness is a type of behavior that is critical for creating and maintaining positive relationships at work and facilitating team functioning. Therefore, promoting assertiveness helps to improve work performance, and facilitate safe behavior and critical decision-making [179].

Analysis (RS.AN):

"Analysis is conducted to ensure effective response and support recovery activities."

Human Factors:

- Communication. To do the response activities you need a strong feedback system to see if the desired results have been achieved [127]. In-person interviews or online questionnaires are good tools for doing this.

Mitigation (RS.MI):

"Activities are performed to prevent expansion of an event, mitigate its effects, and resolve the incident."

Human Factors:

- Knowledge. Countermeasures and installed security procedures are used to reduce the dangers. Cyber-Hygiene can help the organization to mitigate the risk [150].
- Distraction. During mitigation, an employee may perform activities incorrectly due to environmental or social factors [159].
- Awareness. To secure and defend today's company networks from diverse cyber-attacks, awareness is essential. Network administrators and security analysts need to know precisely what happened in the network, why it happened, and what steps or countermeasures need to be made as soon as possible to mitigate any potential effects [208].

Improvements (RS.IM):

"Organizational response activities are improved by incorporating lessons learned from current and previous detection/response activities."

Human Factors:

- Resources and Norms. To strengthen cybersecurity resilience, the organization's framework should include an incident review and learning phase as an official standard. The Lessons Learned section of the cybersecurity incident response process is frequently overlooked, resulting in wasted opportunities that may have helped teams develop, detect critical trends, and enhance their security [209].

Table 18 - NIST function RESPOND & Human Factors

Function	Category	Human Factors	Subcategory	Informative References
RESPOND (RS)	**Response Planning (RS.RP):** Response processes and procedures are executed and maintained, to ensure response to detected cybersecurity incidents.	**Distraction:** If workers experiment inadequate work environment and heavy workload distraction errors may occur when conducting and sustaining response procedures **Fatigue:** Involuntary errors caused by fatigue may occur during response procedures and maintenance of ongoing processes if the employee has a heavy task and a long shift **Knowledge and Awareness:** Stakeholders must be prepared for and respond to security threats by implementing stringent yet multi-faceted cybersecurity measures such as security awareness and incident response training **Pressure:** Employees may experience Time Pressure because there is a time limit on the response tasks that must be carried out **Complacency:** During response process people's seeming complacency is due to	**RS.RP-1:** Response plan is executed during or after an incident	**CIS CSC** 19 **COBIT 5** APO12.06, BAI01.10 **ISA 62443-2-1:2009** 4.3.4.5.1 **ISO/IEC 27001:2013** A.16.1.5 **NIST SP 800-53 Rev. 4** CP-2, CP-10, IR-4, IR-8

Function	Category	Human Factors	Subcategory	Informative References
RESPOND (RS)		the lack of a significant and damaging cyber incident that they had dealing with or because they shift away from self-responsibility during work, sure that others have taken care of the user's system's security without the need for personal involvement		
	Communications (RS.CO): Response activities are coordinated with internal and external stakeholders (e.g. external support from law enforcement agencies).	**Communication:** when in contact with several people, the required soft skills are very important so that so that interpersonal conflict does not sour **Teamwork:** during team communication, the employees must show respect for the other people to improve the teamwork **Fatigue:** Communication with various entities can cause cognitive fatigue and tiredness **Assertiveness:** In conversations between employees and employers, managers and supervisees, and other parties, assertiveness is crucial **Stress:** If two or more interacting people can't	**RS.CO-1:** Personnel know their roles and order of operations when a response is needed	**CIS CSC** 19 **COBIT 5** EDM03.02, APO01.02, APO12.03 **ISA 62443-2-1:2009** 4.3.4.5.2, 4.3.4.5.3, 4.3.4.5.4 **ISO/IEC 27001:2013** A.6.1.1, A.7.2.2, A.16.1.1 **NIST SP 800-53 Rev. 4** CP-2, CP-3, IR-3, IR-8

Function	Category	Human Factors	Subcategory	Informative References
RESPOND (RS)		get along, don't agree, or are otherwise different, there may be interpersonal conflicts that generate general stress	**RS.CO-2:** Incidents are reported consistent with established criteria	**CIS CSC** 19 **COBIT 5** DSS01.03 **ISA 62443-2-1:2009** 4.3.4.5.5 **ISO/IEC 27001:2013** A.6.1.3, A.16.1.2 **NIST SP 800-53 Rev. 4** AU-6, IR-6, IR-8
			RS.CO-3: Information is shared consistent with response plans	**CIS CSC** 19 **COBIT 5** DSS03.04 **ISA 62443-2-1:2009** 4.3.4.5.2 **ISO/IEC 27001:2013** A.16.1.2, Clause 7.4, Clause 16.1.2 **NIST SP 800-53 Rev. 4** CA-2, CA-7, CP-2, IR-4, IR-8, PE-6, RA-5, SI-4
			RS.CO-4: Coordination with stakeholders occurs consistent with response plans	**CIS CSC** 19 **COBIT 5** DSS03.04 **ISA 62443-2-1:2009** 4.3.4.5.5 **ISO/IEC 27001:2013** Clause 7.4 **NIST SP 800-53 Rev. 4** CP-2, IR-4, IR-8
			RS.CO-5: Voluntary information sharing occurs with external stakeholders to achieve broader cybersecurity situational awareness	**CIS CSC** 19 **COBIT 5** BAI08.04 **ISO/IEC 27001:2013** A.6.1.4 **NIST SP 800-53 Rev. 4** SI-5, PM-15
	Analysis (RS.AN): Analysis is conducted to ensure effective response and support recovery activities.	**Communication:** To do the response activities you need a strong feedback system to see if the desired results have been achieved	**RS.AN-1:** Notifications from detection systems are investigated	**CIS CSC** 4, 6, 8, 19 **COBIT 5** DSS02.04, DSS02.07 **ISA 62443-2-1:2009** 4.3.4.5.6, 4.3.4.5.7, 4.3.4.5.8 **ISA62443-3-3:2013** SR 6.1 **ISO/IEC 27001:2013** A.12.4.1, A.12.4.3, A.16.1.5 **NIST SP 800-53 Rev. 4** AU-6, CA-7, IR-4, IR-5, PE-6, SI-4

Function	Category	Human Factors	Subcategory	Informative References
RESPOND (RS)			**RS.AN-2:** The impact of the incident is understood	**COBIT 5** DSS02.02 **ISA 62443-2-1:2009** 4.3.4.5.6, 4.3.4.5.7, 4.3.4.5.8 **ISO/IEC 27001:2013** A.16.1.4, A.16.1.6 **NIST SP 800-53 Rev. 4** CP-2, IR-4
			RS.AN-3: Forensics are performed	**COBIT 5** APO12.06, DSS03.02, DSS05.07 **ISA 62443-3-3:2013** SR 2.8, SR 2.9, SR 2.10, SR 2.11, SR 2.12, SR 3.9, SR 6.1 **ISO/IEC 27001:2013** A.16.1.7 **NIST SP 800-53 Rev. 4** AU-7, IR-4
			RS.AN-4: Incidents are categorized consistent with response plans	**CIS CSC** 19 **COBIT 5** DSS02.02 **ISA 62443-2-1:2009** 4.3.4.5.6 **ISO/IEC 27001:2013** A.16.1.4 **NIST SP 800-53 Rev. 4** CP-2, IR-4, IR-5, IR-8
			RS.AN-5: Processes are established to receive, analyze and respond to vulnerabilities disclosed to the organization from internal and external sources (e.g. internal testing, security bulletins, or security researchers)	**CIS CSC** 4, 19 **COBIT 5** EDM03.02, DSS05.07 **NIST SP 800-53 Rev. 4** SI-5, PM-15
	Mitigation (RS.MI): Activities are performed to prevent expansion of an event, mitigate its effects, and resolve the incident.	**Knowledge:** Countermeasures and installed security procedures are used to reduce the dangers **Distraction:** During mitigation, an employee may perform activities incorrectly due to environmental or social factors	**RS.MI-1:** Incidents are contained	**CIS CSC** 19 **COBIT 5** APO12.06 **ISA 62443-2-1:2009** 4.3.4.5.6 **ISA 62443-3-3:2013** SR 5.1, SR 5.2, SR 5.4 **ISO/IEC 27001:2013** A.12.2.1, A.16.1.5

Function	Category	Human Factors	Subcategory	Informative References
RESPOND (RS)		**Awareness:** to secure and defend today's company networks from diverse cyber-attacks, awareness is essential. Network administrators and security analysts need to know precisely what happened in the network, why it happened, and what steps or countermeasures need to be made as soon as possible to mitigate any potential effects	**RS.MI-2:** Incidents are mitigated	**NIST SP 800-53 Rev. 4** IR-4
				CIS CSC 4, 19 **COBIT 5** APO12.06 **ISA 62443-2-1:2009** 4.3.4.5.6, 4.3.4.5.10 **ISO/IEC 27001:2013** A.12.2.1, A.16.1.5 **NIST SP 800-53 Rev. 4** IR-4
			RS.MI-3: Newly identified vulnerabilities are mitigated or documented as accepted risks	**CIS CSC** 4 **COBIT 5** APO12.06 **ISO/IEC 27001:2013** A.12.6.1 **NIST SP 800-53 Rev. 4** CA-7, RA-3, RA-5
	Improvements (RS.IM): Organizational response activities are improved by incorporating lessons learned from current and previous detection/response activities.	**Resources and Norms:** To strengthen cybersecurity resilience, the organization's framework should include an incident review and learning phase as an official standard.	**RS.IM-1:** Response plans incorporate lessons learned	**COBIT 5** BAI01.13 **ISA 62443-2-1:2009** 4.3.4.5.10, 4.4.3.4 **ISO/IEC 27001:2013** A.16.1.6, Clause 10 **NIST SP 800-53 Rev. 4** CP-2, IR-4, IR-8
			RS.IM-2: Response strategies are updated	**COBIT 5** BAI01.13, DSS04.08 **ISO/IEC 27001:2013** A.16.1.6, Clause 10 **NIST SP 800-53 Rev. 4** CP-2, IR-4, IR-8

5. RECOVER

Develop and implement appropriate activities to maintain plans for resilience and to restore any capabilities or services that were impaired due to a cybersecurity incident.

Recovery Planning (RC.RP):

"Recovery processes and procedures are executed and maintained to ensure restoration of systems or assets affected by cybersecurity incidents."

Human Factors:

- Norms. Establishing clear procedures for incident management is a complex undertaking, and although they are tailored to an organization's mission, size, structure, and functions, they generally contain common elements: (1) defining the scope; (2) defining cybersecurity events and incidents, personnel roles and responsibilities, levels of authority for the response, reporting requirements, requirements and guidelines for external communications, information sharing, and procedures for evaluating performance [210].
- Knowledge. Organizations should use a combination of exercises and tests for recovery capability validation. Recovery exercises and tests should be formally implemented at a frequency that makes sense for the organization, and the results should be recorded to help inform organizational cybersecurity activities. Organizations should set realistic objectives, with specific roles and responsibilities, for exercising and testing recovery capabilities to verify their ability to adequately manage cybersecurity risk [211].
- Distraction. By studying human error based on running multiple processes simultaneously, it was realized that the demand for multitasking distracts workers and encourages cyberattacks [97].

Improvements (RC.IM):

"Recovery planning and processes are improved by incorporating lessons learned into future activities."

Human Factors:

- Resources and Norms. The framework adopted by the organization should establish as an official norm an incident review and learning phase to improve cybersecurity resilience. It is important that Lessons Learned related to recovery plans are not overlooked. Such information is important for improving security and recovery processes [209].

Communications (RC.CO):

"Restoration activities are coordinated with internal and external parties (e.g. coordinating centers, Internet Service Providers, owners of attacking systems, victims, other CSIRTs, and vendors)."

Human Factors:

- Communication and Assertiveness. Especially after experiencing an attack, communicating promptly is critical. One must effectively communicate to the team the activities that are part of recovery planning and reassure external stakeholders. Those who score high in assertiveness are likely to speak up, take the lead, and direct the activities of others. In the context of cybersecurity, a lack of initiative can result in an inability to take responsibility or make an important decision to resolve a cyber attack. In addition, appropriate communication channels must be chosen to communicate a possible attack [210]

Table 19 - NIST function RECOVER & Human Factors

Function	Category	Human Factors	Subcategory	Informative References
RECOVER (RC)	**Recovery Planning (RC.RP):** Recovery processes and procedures are executed and maintained to ensure restoration of systems or assets affected by cybersecurity incidents.	**Knowledge:** Organizations should use a combination of exercises and tests for recovery capability validation **Distraction:** it was studied that the demand for multitasking (e.g. running multiple processes simultaneously) distracts workers and encourages cyberattacks **Norms:** Establishing clear procedures for handling incidents.	**RC.RP-1:** Recovery plan is executed during or after a cybersecurity incident	**CIS CSC** 10 **COBIT 5** APO12.06, DSS02.05, DSS03.04 **ISO/IEC 27001:2013** A.16.1.5 **NIST SP 800-53 Rev. 4** CP-10, IR-4, IR-8

Function	Category	Human Factors	Subcategory	Informative References
RECOVER (RC)	**Improvements (RC.IM):** Recovery planning and processes are improved by incorporating lessons learned into future activities.	**Resources and Norms:** The framework adopted by the organization should establish as an official norm an incident mt review and learning phase to improve cybersecurity resilience.	**RC.IM-1:** Recovery plans incorporate lessons learned	**COBIT 5** APO12.06, BAI05.07, DSS04.08 **ISA 62443-2-1:2009** 4.4.3.4 **ISO/IEC 27001:2013** A.16.1.6, Clause 10 **NIST SP 800-53 Rev. 4** CP-2, IR-4, IR-8
			RC.IM-2: Recovery strategies are updated	**COBIT 5** APO12.06, BAI07.08 **ISO/IEC 27001:2013** A.16.1.6, Clause 10 **NIST SP 800-53 Rev. 4** CP-2, IR-4, IR-8
	Communications (RC.CO): Restoration activities are coordinated with internal and external parties (e.g. coordinating centers, Internet Service Providers, owners of attacking systems, victims, other CSIRTs, and vendors).	**Communication & Assertiveness:** Especially after experiencing an attack, communicating promptly is critical. One must effectively communicate to the team the activities that are part of recovery planning and reassure external stakeholders.	**RC.CO-1:** Public relations are managed	**COBIT 5** EDM03.02 **ISO/IEC 27001:2013** A.6.1.4, Clause 7.4
			RC.CO-2: Reputation is repaired after an incident	**COBIT 5** MEA03.02 **ISO/IEC 27001:2013** Clause 7.4
			RC.CO-3: Recovery activities are communicated to internal and external stakeholders as well as executive and management teams	**COBIT 5** APO12.06 **ISO/IEC 27001:2013** Clause 7.4 **NIST SP 800-53 Rev. 4** CP-2, IR-4

4.5. Results and discussions

The IDENTIFY function is related to: (i) understanding the resources that support crucial business operations; (ii) the business environment; (iii) the associated risks. When discussing resources (i), it means both human and technical resources. In this context, the factor "Resources" plays a crucial role. The organization must allocate enough budget to acquire all the necessary resources and maintain them for effective cyber resilience. If resources are insufficient, employees may suffer from "stress" [182], [183]. When it comes to human resources, the "communication" factor plays a key role. Clearly defining and assigning responsibilities that match the capabilities of individuals helps to create a climate of mutual respect between management and employees, which is critical to achieving strategic goals. To achieve these goals, it is necessary to have the "Knowledge" of how to set priorities and use resources optimally. The business environment (ii) includes the business environment and governance. Organizations depend critically on top management's ability to create a business environment in which everyone has "awareness" of cybersecurity issues [192]. Spreading awareness does not mean establishing rigid rules that are difficult to enforce. Much of the literature shows that applying very rigid policies is a major cause of "stress," "fatigue," and "pressure" in the cyber environment. A good way is to organize awareness campaigns and create an environment where employees are motivated and stimulated to increase their "Knowledge" and are free to communicate ("Communication") without shame and fear of consequences. The associated risks (iii) are also related to the supply chain. In the IDENTIFY phase, this risk assessment phase is critical. It is critically important that all members of the supply chain are aware of cyber risks, but more importantly that they have adopted common "standards" for managing operations and data. To be monitored at this stage is the factor of "complacency" [134], [136].

The PROTECT function supports the ability to limit or contain the impact of a potential cybersecurity event. Through the PROTECT function, measures are put in place to protect people (1), data (2), and assets (3). People (1) are protected through identity management, authentication and access control, awareness, and training. The organization must carefully regulate access points to prevent unauthorized access to authorized activities and transactions. It is necessary to communicate ("Communication") the authorized employees and monitor the multiple access points. In addition, one should establish "Norms" that are effective but not difficult to enforce. Often if the rules are repetitive, such as frequent password change requests or constant requests for authorizations, the employee may experience "Fatigue," so-called "authentication fatigue" [212]. In addition, due to the repetitiveness of actions, he or she may make "Distraction" errors. Protecting access points is critical, but protection works if employees are aware ("Awareness") of cyber threats and trained to know ("Knowledge") their impact and likelihood of occurrence.

The topic of training returns in many of the functions. This, to be effective, must be structured to be stimulating and instructive for the employee. The employee is most likely to feel "fatigued" or "stressed" if the training involves a heavy workload, or shows complacency towards it [112], [128], [133]. Information and data (2) must be managed in a manner consistent with the organization's risk strategy to protect the confidentiality, integrity, and availability of information. This means adopting security "communication" principles to protect data in transit [202].

In addition, organizations need to structure protection systems for the user. It is suggested that the design consider characteristics of individuals such as distraction, stress, fatigue (in all its categories described by [213]), and the need to go fast at the expense of safety [177], [183], [201]. In addition, in setting norms, the organization must ensure that people do not feel too excluded, trained, constrained, and controlled to comply with them, increasing resistance [117].

As for the DETECT function, again the human factors "Knowledge" and "Awareness" play a key role. When a detection process has to start, one should know how to assess the threats and their implications. In addition to these is the importance of the human factor "Teamwork." Especially when it comes to the detect function, the interaction between teams could lead to better cyber defense performance [204].

The most critical human factors that organizations should pay attention to are the human factors "Pressure" and "Distraction." "Pressure" for workers directly involved in processes to detect attacks. Misjudgments caused by excessive work pressure can have damaging effects on business and personal data [173]. Therefore, organizations must act in a way that does not overburden employees with pressure when assigning tasks and responsibilities. Finally, distraction-related to continuous security detection and monitoring processes should be mentioned. When employees place too much trust in cybersecurity devices, this can reduce control and commit "distraction" errors due to the repetitiveness of the actions they perform [136].

Analyzing the last two remaining functions, "RESPONDING" AND "RECOVERY," the human factors "Knowledge" and "Awareness," "Stress," and "Pressure" need to be mentioned again. The first two are among the most analyzed in the literature. They are a fundamental requirement not only for the NIST categories but for cyber resilience in general. No process or procedure can be implemented without a thorough understanding of operations and situational and contextual awareness. The connection between these functions and "Stress" and "Pressure" is also intuitive. Stress management is a primary consideration in any incident response team. Incident response work is technical, laborious, and difficult and often must be done quickly to reduce damage and safeguard the existence of the organization.

However, there are other factors found to be important in the analysis of the two functions. Among the most cited in the literature are "communication" and "norms." When an organization is hit by an attack, the impact is likely to affect not only the organization itself but also other stakeholders. This means that the organization must communicate with different entities during the response and recovery phases. Soft communication skills are essential at this stage [130], [209]. In addition, in order to respond and recover effectively, a strong feedback system is needed to check whether the desired outcomes have been achieved [127].

Moreover, it has emerged how lessons learned related to the cybersecurity incident response process are often overlooked. The result is that organizations miss opportunities to help teams mature, identify important trends, and improve security.

However, these processes involving learning and improvement only work if incidents are reported.
Nevertheless, reporting is often neglected. At the individual level, the reason is shame and fear of consequences. At the corporate level the need to protect the company image [117]. It is important to understand the motivations and benefits of reporting and to create a friendly and collaborative work environment where communication does not become an obstacle to cyber resilience. Finally, the last widely cited factor is "Assertiveness." After experiencing an attack, communicating promptly is critical. As mentioned above, communicating effectively with one's team about activities that are part of the response and recovery planning and reassuring external stakeholders about the situation should be two priorities. Those who score high in assertiveness are likely to speak up, take the lead, and direct the activities of others. In the context of cybersecurity, a lack of initiative can result in an inability to take responsibility or make an important decision to resolve a cyber attack. Assertiveness is a key behavior for creating and maintaining positive relationships at work and facilitating team functioning and post-incident recovery.

The following framework, graphically summarizes the result of this latest analysis, emphasizing the most impactful human factors in the NIST functions.

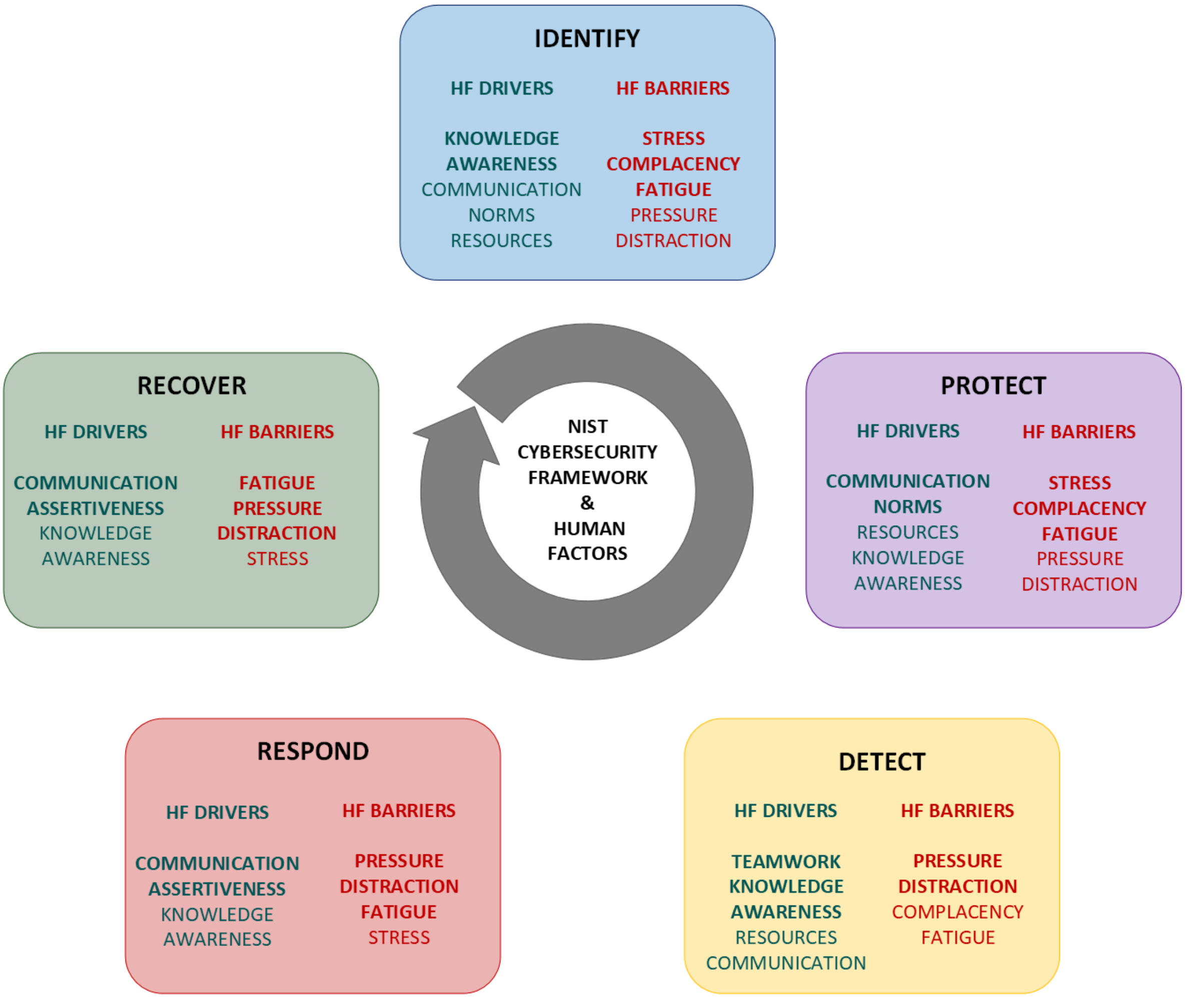


Figure 13 - NIST Cybersecurity Framework & Human Factors. In bold most relevant factors per function

5. Discussing the effectiveness of organizational cybersecurity outsourcing practices

Over the past few years, organizational cybersecurity has become a popular topic of discussion. [214] define organizational cybersecurity broadly as *"the efforts organizations take to protect and defend their information assets, regardless of the form in which those assets exist, from threats internal and external to the organization".* Organizations of all sizes are now looking to implement measures to protect themselves from cyberattacks. While no one solution will work for every organization, implementing a combination of technologies seemed the most effective way to reduce risk. Proper technology can help prevent most attacks, detect vulnerabilities, and mitigate cyber risks [215]. However, it is now clear that cybersecurity requires more than just the latest technology [191]. To decrease cyber risk, all members of the organization must act. Working to identify and reduce risks, implementing rules and procedures, and educating staff are all part of proper security planning [216] involved in cybersecurity management. An important context-based variable to consider when dealing with organizational cybersecurity is represented by the choice of outsourcing (or not) cybersecurity management [217]. This can be due to the specific skills owned by the organization [218], the organizational dimension [155], [219], the cybersecurity budget [220], and other contextual variables. As extensively demonstrated in other managerial research streams, there are significant differences in the effectiveness of managing internally or outsourcing organizational and business processes, and cybersecurity processes are no exception [221], [222]. While outsourcing has led organizations to achieve goals of reducing costs, simplifying operations, improving productivity, and enhancing customer service [223], today organizations are questioning whether this strategy is increasing or reducing threats and risks. Organizations around the world are looking to improve and grow their business by focusing on their core activities, which has led them to increasingly rely on external staff to manage specific aspects of their organizational processes. It is in this context that IT outsourcing (ITO) has seen a significant increase, confirmed by a proliferation of contributions that stressed the importance of defining ITO models to help organizations in the decision-making process for choosing to outsource their services to third parties. However, while ITO continues to be popular for its ability to make enterprises more agile and cost-effective, the associated cybersecurity concerns have been growing and taking on a more urgent priority. According to [224], the benefits of the expanded use of outsourced services may be impeded by the increased potential cyber risk exposures that these services create for the companies who acquire them. Companies that use outsourcing services may be under-protected against these hazards or may not even be aware that they exist. On the other side [225], [226] suggest that companies that outsource cybersecurity would benefit from high-quality software and highly skilled cybersecurity experts which lead to a reduction in cyber risks. Although in the last years there has been a proliferation of contributions to guide organizations in the management of cybersecurity [227], [228], there is still a research gap in studying the differences concerning the effectiveness of cybersecurity practices in the case of in-house managed or outsourced cybersecurity processes. Most existing research studies information security outsourcing as an operational IT decision [229] or as an attack detection and secure estimation problem [230], [231]. Very few studies investigated strategic organizational decisions and management processes involved in the decision-making process of outsourcing information security. With this research, the author wants to contribute to the debate on the decision-making process for choosing to outsource cybersecurity by stressing the importance of also considering managerial aspects and variables related to the organizational context. In the following paragraphs will be presented a summary of the prior literature related to the definition and utilization of ITO and cybersecurity outsourcing. Then a paper titled "The Effectiveness of Outsourcing Cybersecurity Practices: A Study of the Italian Context. published in the Proceedings of the Future Technologies Conference will present the results of a quantitative study that surveyed 153 experts in cybersecurity on the effectiveness of cybersecurity practices.

5.1. IT Outsourcing

Outsourcing is defined as a decision-making process where the management of the organization must decide whether they should keep a specific activity in-house or buy it from an external subcontractor [232]. Multiple definitions have been proposed for ITO. Among the most common are [233] which define ITO as the use of third-party service providers to effectively deliver IT-enabled business processes, application services and infrastructure solutions for business outcomes; [234], [235] integrate the definition mentioning also connectivity, development, and maintenance of both hardware and software; a fourth definition is given by [236], [237] which describes ITO as "a decision taken by an organization to contract out or sell the organization's IT assets, people, and/or activities to a third party vendor, who in exchange provides and manages assets and services for monetary returns over an agreed period [237].
More specifically [238] in their state of the art on ITO states that activities that are outsourced can be handled by four different types of agreements: (i) General outsourcing, which can be done by selecting an area of its information system functions according to a strategic plan or identifying an area which could gain value if outsourced. (ii) Transitional outsourcing in which the organization replaces its technical platform with another one through a third party. (iii) Business process outsourcing in which a third party runs all the functions of the business. (iv) Business benefits contracting in which an organization subcontracts with a third party defining

the business benefits to be achieved in a period. In this thesis, ITO will be considered a generic term that also covers the evolutions of approaches such as net-sourcing, cloud sourcing, offshoring, and quasi-outsourcing. Net-sourcing is the practice of renting or “paying as you use” access to a centrally managed business application, made available to multiple users from a shared facility over networks [239]. Cloud Sourcing enables organizations to purchase IT resources and capabilities from another organization as a service. In this case, organizations can choose to outsource all or part of their IT services to the cloud to run applications, databases, and servers on a virtualized infrastructure [240]. Offshoring or outsourcing involves a contract with an IT service provider outside of its home country with privacy, property, security, and regulatory compliance implications [241], [242]. Quasi-outsourcing involves the creation of a subsidiary and transferring certain business functions to it, keeping total or partial ownership of the new, independently managed company [243].

According to [244] the main driver for which a company chooses to outsource part, or all of its IT activities is cost savings, but it is necessary to consider further aspects such as: the greater speed of development in implementing key functions compared to a more expensive and slower internal activity; the operational flexibility needed by a fast-growing company to maintain high demand; the highly specialist skills required by the IT context. Outsourcing also enables organizations to better focus on their core business while also improving the service they offer [245]. and reducing the risks associated with compliance with privacy-related issues [246]. Finally, several variables need to be taken into consideration when evaluating whether to outsource processes and which form of outsourcing is best suited to the business. [171] discusses the link between organizational size and the decision to externalize processes. Specifically, their case study details factors that influence the effectiveness of outsourcing for small and medium-sized enterprises. [247] on the other hand, highlights the differences in outsourcing IT processes in the public or private sector, listing the motivations that lead a public company to outsource its IT processes and describing the difficulties of ITO process management in the public sector. Finally, [245] mentions the importance of considering the final customer in decision-making factors, especially in B2C organizations, it is stated how ITO must be an opportunity both to maximize its benefits and to ensure the satisfaction of its internal and external customers.

5.2. Cybersecurity Outsourcing

Within the ITO stream, it is important to highlight the issue of cyber security. On one side cyber security has become one of the top priorities for IT managers in both public and private sectors with large and small organizations facing increasingly sophisticated methods of attacks. On the other side, it has also drawn the attention of researchers and practitioners to investigate whether cyber security outsourcing, as part of the ITO, would lead to benefits or increased risks for today's industrial contexts [224]. A continuously growing number of companies have approached the paradigm of industry 4.0, by connecting factories to the internet, allowing business functions to communicate in real time, to increase efficiency and effectiveness. However, in these hyper-connected industrial contexts, cybersecurity issues represent one of the most relevant challenges and barriers to efficiency. Within Industry 4.0, cybersecurity plays a key role in preventing companies from losing competitiveness and it is perceived by managers to be both a top priority and one of the most critical shortcomings [248]. A cybersecurity program for a highly connected industry is complex. It has several components from strategy to architecture, engineering to operations to be taken into consideration [249]. Organizations can manage in-house their cybersecurity process or pay a service provider [217]. The strategic outsourcing of cybersecurity functions explicitly (or implicitly) assumes that organizations transfer the responsibility to cybersecurity providers. These organizations' risk profile changes and becomes a combination of their risks and a subset of their cybersecurity provider risks [224]. Since cybersecurity risk is never totally transferred, cybersecurity practices are compulsory in both contexts (in-house and outsourcing) even if it is plausible that one could be a better choice in specific circumstances. A branch of studies has focused on information security processes, patching policies, contractual issues, and attack control and monitoring software, rather than investigating the strategic decisions, cybersecurity management practices, and IT governance arrangements associated with the decision to outsource cybersecurity [229]. Another branch of studies has focused on the effectiveness of cybersecurity practices from an organizational point of view, considering those aspects that do not involve only technology but are also considered fundamental to the analysis of business processes, the engagement of people, and the interaction between man and machine within the decision-making process [47]. In conclusion, previous studies have focused on the benefits and risks of both solutions, however, there is still a research gap in studying the differences concerning the effectiveness of organizational cybersecurity practices when they are outsourced.

We will refer to the concept of effectiveness as the degree to which a practice is successful in producing the desired result: the successful information security of the company.

The study aims to assess managers' perceptions of the effectiveness of selected organizational cybersecurity practices in protecting their corporate systems by going to evaluate cases of in-house or outsourced cybersecurity management.

Specifically, the quantitative study conducted involved 153 managers in assessing the effectiveness of a selected group of practices related to organizational cybersecurity.

In fact, for this phase of the research, we focused on those practices that according to the NIST framework most relate to the interaction with the characteristics of the employees involved in cybersecurity processes.

The focus was on aspects such as:

- The management of cybersecurity process control, particularly the presence of disciplinary processes and sanctions for information security data breaches.
- The procedures that regulate physical access to IT resources.
- The practices to ensure the protection of log information, specifically the audit and log record should be documented, implemented, and revised according to the policy.
- The use of lessons learned in response plans.
- The use of communication protocols to communicate with internal and external stakeholders.

The results answered RQ3 showing that there is no difference in the effectiveness of procedures that regulate physical access to IT resources or communication protocols, going to confirm the current trend of a balanced partition between those organizations that consider it less hazardous to keep such practices in-house and those that prefer to outsource them convinced that they would find out more trained people.

Interesting, however, is the result that emerged on the topic of sanctions and disciplinary processes. What emerged in the human factors analysis in chapter 4 showed that multiple situations lead a user to commit an information security breach. There are motivations related to a desire for revenge, frustration due to fatigue and pressure, distraction, or complacency toward cyber security practices.

NIST states that organizations should establish a formal and communicated disciplinary process to take action against employees who do not comply with the company policy. What emerges is how deterrence actions and severity of penalty seem to be more effective when people from outside the organization handle them. Individuals from outside the organization may be less emotionally involved making such processes more effective and increasing the intention of employees to comply with organizational information security policies.

Regarding data protection and recording lessons learned, it was found that managers find it more effective to keep their management in-house. Regarding data, the result is in line with the broader literature on cloud-based data management. There is a general mistrust and concern about outsourcing large volumes of data. This has led organizations to develop and train in-house IT teams responsible for such management.

Finally, the theme of lessons learned. Organizations are still disinclined to record IT incidents and even more so to outsource this activity [250]. There is still a strong view that disseminating such information is inconvenient for the organization's reputation. However, the study showed how organizations have begun to understand the importance of incident analysis and recording lessons learned, valuing this practice as more effective when it is done internally within the organization. The reasons for a higher effectiveness value can be attributed, according to managers, to incentives to withhold information about cyber attacks, especially when the occurrence of the cyber attack and the damage caused is uncertain.

For a detail of the survey conducted and the results collected for the practices mentioned above, please refer to the follow-up article presented at the Future Technologies Conference SAI 2021. The presented article is intended to be a starting point that analyzes specific cybersecurity practices and lays the foundation for a larger study in the future aimed at analyzing more aspects related to cybersecurity outsourcing.

5.3. Appended Paper 2: The Effectiveness of Outsourcing Cybersecurity Practices: A Study of the Italian Context

The effectiveness of outsourcing cybersecurity practices: a study of the Italian context

Alessandro Annarelli[1], Silvia Colabianchi[1], Fabio Nonino[1], Giulia Palombi[1]

[1] Department of Computer, Control and Management Engineering, Sapienza University of Rome, Rome, Italy
alessandro.annarelli@uniroma1.it,
silvia.colabianchi@uniroma1.it, fabio.nonino@uniroma1.it,
giulia.palombi@uniroma1.it

Abstract. The increasing number of cyber-attacks requires an organizational awareness about the disruptive effects of fraud attempts and acts of vandalism on business continuity and, sometimes, on company survival. The context influences the way companies use and adapt these theories in practice, so we consider in this study differences in the effectiveness of cybersecurity best practices between organizations that manage internally or outsource the cybersecurity processes. We conducted a study involving 153 managers experts in cybersecurity who responded to a survey on the effectiveness of NIST procedures. Results revealed significant differences in the effectiveness of managing cybersecurity in-house or outsource it. Specifically, major differences can be observed in the variables related to the use of disciplinary processes, the protection of log information, and the use of lessons learned to improve recovery plans. These differences provide further insights for cybersecurity management literature and a practical instrument for organizations willing to adapt their cyber processes to their organizational context.

Keywords: NIST framework, cyber risk, confirmatory study, internalization, externalization.

1 Introduction

In the last years, the increasing importance and criticality of cyberspace due to its role in the digitalization process of business and private and public services calls for the need to design effective cybersecurity practices. Many approaches [1], [2] and frameworks [3]–[5] for cybersecurity have been proposed in the literature. One of the fundamental and internationally recognized is the NIST framework [3]. NIST offers guidance and provides guidelines, best practices, and standards for research and applications in cybersecurity risk management. Nevertheless, like other frameworks, it contains general guidelines that must be carefully adapted to the target organizational context [6], [7]. An important context-based variable to consider is represented by the outsourcing (or not) of cybersecurity management [8]. This can be due to the specific skills own by the organization [9], to the organizational dimension [10], [11], to the cybersecurity budget [12], and other contextual variables. As extensively demonstrated in other managerial research streams, there are significant differences in the effectiveness of managing internally or outsourcing organizational and business processes and cybersecurity processes are no exception [13], [14]. In the last years, there has been a proliferation of contributions from literature and international standards in order to guide organizations in the management of cybersecurity, but there is still a research gap in studying the differences concerning the effectiveness of cybersecurity practices in the case of in-house managed or outsourced cybersecurity processes.

We conducted a survey involving cybersecurity experts and we tested statistically some hypotheses. We found interesting effects of the strategic choice of cybersecurity outsourcing: the major effectiveness of disciplinary processes in case of non-compliance with company policies for the organizations and of protection of log information. Fundamental is, instead, the role of the lessons learned for organizations that decide to manage internally the cybersecurity process.

2 Theoretical background and hypotheses

Organizations can manage in-house their cybersecurity process or pay a service provider [8]. The strategic outsourcing of cybersecurity functions explicitly (or implicitly) assumes that organizations transfer the responsibility to cybersecurity providers. These organizations' risk profile changes and becomes a combination of their risks and a subset of their cybersecurity provider risks [15]. Since cybersecurity risk is never totally transferred, cybersecurity practices are compulsory in both contexts (in-house and outsourcing) even if it is plausible that one could be a better choice in specific circumstances. The NIST framework [3] suggests a large number of cybersecurity practices for organizations organized in five functions: identify, protect, detect, respond, and recover. Each function is articulated in categories and subcategories. For each subcategory, a list of international standards' practices to implement is provided. In the protection and recovery categories, we identified some practices that, according to literature, seem to have different effectiveness for organizations managing in-house or outsourcing cybersecurity processes.

2.1 Disciplinary process

Among practices for the *Information Protection Processes and Procedures*, NIST reports the importance of integrating human resources practices with cybersecurity [3]and, according to [16] "*there shall be a formal and communicated disciplinary process in place to take action against employees who have committed an information security breach.*" [16, p. 11] . Research has proven that employees' perceived severity of cyber-attack might significantly affect security concern [17] and their perception of why and how disciplinary processes can improve the overall information systems' effectiveness [18], [19]. On the other hand, employees' perceived severity is not a strong predictor of their cybersecurity protective intention due to a lack of people's awareness of the real threats posed by information security incidents [20]. Employees' sense of vulnerability is a key element in the context of cybersecurity threats appraisal. Indeed, an employee, in perceiving high vulnerability risk to his/her organization's information systems, will more likely take protective actions. Furthermore, employees' perceived vulnerability in the cyber-attack incident positively motivates them to comply with cybersecurity policies [21], [22] and their perception of security importance and social influence plays a key role in ensuring commitment to cybersecurity issues [17]. Recently it has been demonstrated that only perceived vulnerability of the individual and the perceived severity of consequences for the organization affect perceived threats, influencing protection motivation [23]. Following these important results, we want to understand what can happen when these incentives are missing, i.e. when cybersecurity is not internally managed, and there is a lack of social pressure (or no pressure at all) to comply with cybersecurity standards. Furthermore, outsourcing relationships are often hindered by a series of trust barriers affecting both the vendor and client side. As highlighted by [24], on one hand, the client is characterized by concerns that involve security controls (put in place by the vendor) as well as the security and privacy of data; on the other hand, among concerns on the vendor's side, the competency and reliability of the client play a central role. In these cases, companies must put in place stronger disciplinary processes, relying on penalties like sanctions and detention [17]. Therefore, we formulate the following hypothesis.

Hypothesis 1. There are significant differences in the effectiveness of the use of disciplinary processes, for those who do not comply with the company policy, in organizations managing cybersecurity internally and outsourcing it

2.2 Procedures that regulate physical access to IT resource

The *Identity Management, Authentication and Access Control* practice suggests that *"access to physical and logical assets and associated facilities must be limited to authorized users, processes, and devices, and is managed consistent with the assessed risk of unauthorized access to authorized activities and transactions"* [3, p. 29]. It provides many actions to undertake to ensure security including how to authorize user access, to prevent unauthorized access to systems and services and how to make users accountable for safeguarding their authentication information [16].

Following these recommendations, it clearly emerges how ensuring the protection from cyber threats for a company is also a matter of physical security and access restriction to physical IT resources, e.g. computers and workstation. It is not solely the need of excluding access to external (or unauthorized) figures, but it is also a matter linked to the need of tracing accesses to sensible information: that is why it is vital for companies to define clear procedures (and policies) to this regard [25], [26]. This practice becomes even more important and somewhat delicate if we think that usually an outsourcing relationship poses a series of trust concerns between the client and the vendor/provider [24] and therefore we believe that this practice might be more effective when handled internally rather than externally, since it may cause more serious trust issues if an external provider should put in action measures to restrict/limit physical access to internal resources.

Hypothesis 2. There are significant differences in the effectiveness of the use of procedures that regulate physical access to IT resources in organizations managing cybersecurity internally and outsourcing it

2.3 Practices to ensure protection of log information

Following the need to restrict and trace access to physical resources, it is important to highlight the need to protect log information as well. According to the NIST Framework, in the Protective Technology category, it is vital that *"audit/log records are determined, documented, implemented, and reviewed in accordance with policy."* [3, p. 36]. More in detail, ISO/IEC 27001 contains a complete set of practices concerning Logging and Monitoring: companies must record *"user activities, exceptions, faults and information security events, [...] log information, [...] system administrator and system operator activities"*, protect and review these information [16, p. 16]. The importance of log records is not only determined by the importance and sensible nature of these information per se, but it is also strictly connected to the response and recovery ability of a company when an accident or a cyber-threat might occur. The vendor-client relationships, that occur when companies outsource IT and cybersecurity practices, are characterized by a series of concerns involving reciprocal trust, e.g. the ability of the vendor to comply with clients' security standards, together with the management of sensible information and privacy data [24], [27]. Following these issues, we believe that practices to ensure the protection of log information (and other sensible data as well) might be more effective when handled internally, rather than externally. Hypothesis 3 details and formalize the importance of this factor.

Hypothesis 3. There are significant differences in the effectiveness of the use of practices to ensure the protection of log information in organizations managing cybersecurity internally and outsourcing it

2.4 The use of lessons learned into response plans

The use of lessons learned into response plans is included in the NIST *Improvement* category, Recover phase. According to NIST, *"organizational response activities are improved by incorporating lessons learned from current and previous detection/response activities"* [3, p. 43]. It means that knowledge obtained from analyzing and resolving information security incidents shall be used to reduce the likelihood or impact of future incidents. The organization shall describe and use procedures for the identification, collection, acquisition, and preservation of information, which can serve as evidence [16].

In order to help and support employees in identifying security threats and determining their severity, it is important for companies to align the organizational environment with IT security needs and awareness. [28] stated that the organizational environment should help the employees in managing the trade-off between the consolidation of proven practices and the exploration of new solutions for information protection and threats mitigation, in accordance with the information compliance procedure. To this extent, lessons learned deriving from past incidents play a key role in guiding the company and its employees in defining the best strategies, first of all for what concerns recovery plans, but also in terms of protection from future risks and potential threats [28]. Moreover, according to [29] when outsourcing cybersecurity and IT processes, there are often difficulties for the managed security service provider (MSPP) to perfectly observe and understand accidents from an *ex-post* perspective, and those difficulties pose a serious problem whenever there is a need to carefully translate what happened into a repository of lessons learned for future recovery plans. Following the above reasons, we believe that the incorporation of lessons learned into recovery plans can be a more effective practice if adopted when cybersecurity is managed internally.

Hypothesis 4. There are significant differences in the effectiveness of the use of lessons learned into response plans in organizations managing cybersecurity internally and outsourcing it

2.5 The use of communication protocols to communicate with internal and external stakeholders

The last factor taken into account for this study is communication and communication protocols. The NIST Framework, for what concerns the *Recover phase and Communications* category, reports the need for organizations to communicate the recovery plan and related activities *"to internal and external stakeholders as well as executive and management teams."* [3, p. 44]. Looking at the ISO/IEC 27001 standard (Clause 7.4), there are more details about this aspect: *"the organization shall determine the need for internal and external communications relevant to the information security management system including on what to communicate, when to communicate, with whom to communicate, who shall communicate, and the processes by which communication shall be effected."* [16]. Communication protocols play indeed a central role in many different activities and are considered to be at the core of an effective and efficient organization [30]. This role becomes even more crucial in the field of risk management, following the importance to quickly and effectively interact with stakeholders during and after an incident, to maximize the response and recovery capabilities [31]–[33]. According to [34] relationships with stakeholders might become tense when cyber threats happen, and therefore it is crucial to accurately put in place protocols to effectively manage communications when incidents occur. Furthermore, this might become even more critical when involving third parties (i.e. when outsourcing cybersecurity and IT processes) that should manage these communication protocols as well [34]. Moving from the previous statements, we believe that the use of communication protocols can be a more effective practice if adopted when cybersecurity is managed internally. Therefore, we formulated the last hypothesis as follows.

Hypothesis 5. There are significant differences in the effectiveness of the use of communication protocols to communicate with internal and external stakeholders in organizations managing cybersecurity internally and outsourcing it

3 Method

3.1 Data collection and analysis

We collected data through an online survey over a one-month period in July 2020. We contacted by e-mail 1539 qualified managers recorded in the "List of Qualified Managers and Consulting Companies" published by the Ministry of Economic Development [35]. Two main criteria were adopted in selecting the participants. First, managers should work in the cybersecurity or IT area. Second, they should have experienced at least one year working within cybersecurity. A total of 153 managers completed the online questionnaire (response rate: 10%). We asked participants to fully complete the questionnaire without leaving any blank questions so, for this reason, 36 participants who did not match the criteria were removed.

The questionnaire consists of several parts. The first part was about the background information of the respondents and their organizations. In the second part, it was asked to rate the effectiveness, within the respondent's organization, of the NIST practices proposed in the questions. An ordinal scale was used [36]. The 5 responses and the scores used to conduct the analysis are as follows: (1) not at all; (2) a little; (3) enough; (4) a lot; (5) very much. A Mann-Whitney U test [37] was used to test the differences between the two groups: organizations who manage cybersecurity internally and those who manage it externally. The data satisfied the three assumptions underlying the Mann-Whitney U test: (1) the two samples are random, (2) the two samples are independent, and (3) the scale of measurement is ordinal.

3.2 Measures

Demographic information
The participants were asked to provide individual demographics including their age and gender, as well as their education level, study area, and the number of years they have worked in cybersecurity.

Organization information
The participants were asked to provide information related to their organization: the size of the company; the number of employees involved in cybersecurity; and how long has the organization been investing in cybersecurity. Moreover, it was asked whether their company works business-to-business or business-to-customer and if they manage cybersecurity externally or internally.

NIST practices

For the development of the questions aimed at assessing the effectiveness of cybersecurity practices, functions, categories, and corresponding subcategories presented in the NIST framework were examined. In particular, the ISO/IEC 27001:2013 standard [16] was used - as also suggested in the NIST framework - as the reference standard for the selection and comprehension of the practices to be implemented and the development of the questions. This standard was selected for its international relevance and its widespread use within business organizations in which the respondents work [38], [39]. **Fig. 1** shows the process used to formulate the questions. First, the functions and categories that were described in the NIST framework as inherent to the managerial context were selected. Subsequently, to select the subcategories, the authors identified those that, according to their experience, were most relevant to the managerial context and of greatest interest, according to the literature, for the decision to manage cybersecurity in-house or outsource it.

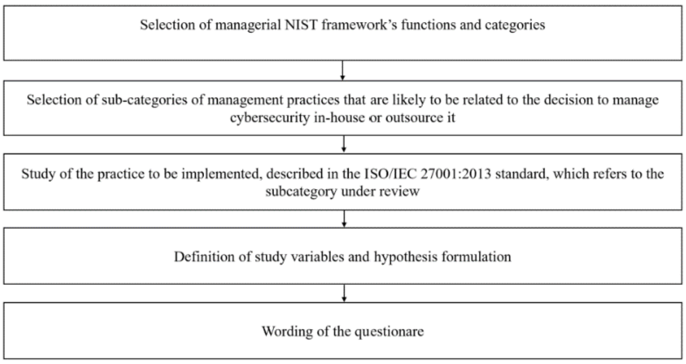

Selection of managerial NIST framework's functions and categories
Selection of sub-categories of management practices that are likely to be related to the decision to manage cybersecurity in-house or outsource it
Study of the practice to be implemented, described in the ISO/IEC 27001:2013 standard, which refers to the subcategory under review
Definition of study variables and hypothesis formulation
Wording of the questionare

Fig. 1. Questionnaire design process

Table 1 below shows the questions of the survey. For each question, the category and subcategory to which it belongs have been included. Moreover, it is reported also the name of the practice and its description provided in the ISO/IEC 27001:2013 standard [16].

Table 1. NIST practices and survey questions

NIST Function and category	**NIST Sub-category**	**Practice**	**Description (ISO 27001:13)**	**Survey Question**
Protect: Information Protection Processes and Procedures	Cybersecurity is included in human resources practices	Disciplinary process for information security breaches	There shall be a formal and communicated disciplinary process in place to act against employees who have committed an information security breach	In terms of Cybersecurity, how critical is that workers who do not comply with the organization's policy are subject to sanctions?

Protect: Identity Management, Authentication and Access Control	Physical access to assets is managed and protected	Procedures that regulate physical access to IT resources	There is the need for companies to define security perimeters to protect areas that contain sensitive or critical information. These areas should also be protected by proper entry controls to ensure that only authorized personnel are allowed access.	How effective are the procedures employees must follow to reduce vulnerabilities in your systems?
Protect: Protective Technology	Audit/log records are determined, documented, implemented, and reviewed in accordance with policy	Protection of log information	Logging facilities and log information shall be protected against tampering and unauthorized access	In terms of Cybersecurity, how crucial is the way your information system keeps track of employee access and activity?
Recover: Improvements	Recovery plans incorporate lessons learned	Lessons learned into recovery plans	Knowledge gained from analyzing and resolving information security incidents shall be used to reduce the likelihood or impact of future incidents	How crucial are the lessons learned at the end of a Cyber incident in preventing future attacks?
Recover: Communications	Recovery activities are communicated to internal and external stakeholders as well as executive and management teams	Communication protocols to internal and external stakeholder	The organization shall determine the need for internal and external communications including on what, when, with whom to communicate; who shall communicate; and the processes by which communication shall be effected.	How crucial is the presence of a protocol on internal and external communications to be followed in case of cyber-attack?

4 Results

4.1 Demographics and organizational variables

Participants were well distributed across age categories both in organizations that manage cybersecurity internally or that outsource it. Approximately 46% of participants were between 41 and 50 years of age. This left approximately 18% in the less than 40 years of age category and 29% in the 51 to 60 age category, and 7% of the participants were 61 years and over. The majority of participants (n=77) had a master's degree from a computer science-related school (n=33) or engineering school (n=53).

Considering the organizations in which participants work, the analysis showed that the dataset was quite balanced across the size of the organization. In addition, despite managing cybersecurity internally or externally, the analysis showed that respondents work more in organizations with fewer than 50 cybersecurity employees and are primarily business to business. **Table 2** below provides some background information about the respondents and the organizations.

Table 2. Background of respondents and the organizations

	In-house	**Outsourcing**
Total responses	66 (56%)	51 (44%)
	%	**%**
Gender		
Male	97%	88%
Female	3%	12%
Age		
Younger than 40	18%	18%
41-50	45%	47%
51-60	29%	29%
61 and above	8%	6%
Educational Level		
High School diploma	15%	16%
Bachelor degree	11%	2%
Master of Science degree	52%	53%
Master of Arts degree	11%	20%
PhD - Doctor of Philosophy	12%	10%
Study Area		
Law	2%	4%
Management	11%	18%
Science	8%	6%
Engineering	47%	43%
Electronics and Telecommunications	2%	2%
Computer Science	30%	25%
Cyber Security	2%	2%
Years in Cyber		
Less than 5	29%	22%
From 6 to 10 years	18%	35%
From 11 to 20 years	38%	27%
from 21 to 30 years	14%	12%
More than 30 years	2%	4%
Size		
Micro	44%	35%
Small	29%	24%
Medium	8%	20%
Large	20%	22%
Number of employees involved in cyber security		
Less than 50	79%	86%
From 51 to 100	15%	10%

More than 100	6%	4%
Business to business / Business to customer		
Business to business	86%	73%
Business to customer	14%	27%
Number of years of investment in cyber security		
Less than 5 years	41%	55%
From 6 to 10 years	20%	14%
From 11 to 20 years	27%	20%
from 21 to 30 years	11%	10%
More than 30 years	2%	2%

4.2 NIST practices

The participants were asked to indicate how effective are managerial NIST practices in their organization in the management of cybersecurity. The results of the analysis are shown in Table 3. The results of the statistical test of difference (Mann-Whitney U) show a statistically significant difference between organizations who manage cybersecurity in-house or that outsource it for practices related to the disciplinary process, protection of log information, and lessons learned after an attack occurred. Specifically, the analysis shows that managers consider the use of disciplinary processes to be more effective in the case of outsourced cybersecurity. On the other hand, the protection of log information and the recording of lessons learned result to be more effective in the case of in-house cybersecurity. The survey responses did not reveal significant differences between the two groups for the other tasks investigated.

Table 3. Mann-Whitney U Test results

Practice	**In-house**		**Outsourcing**		**Mann-Whitney U Test**	**More supportive group**
	Mean	**Std.**	**Mean**	**Std.**		
Disciplinary process for information security breaches	2.83	1.046	3.24	0.992	**0.040***	Outsourcing
Procedures that regulate physical access to IT resources	3.82	0.893	3.86	0.96	0.635	-
Protection of log information	3.98	0.832	3.65	0.82	**0.030***	In-house
Lessons learned into recovery plans	4.38	0.651	4.06	0.705	**0.014***	In-house
Communication protocols to internal and external stakeholder	3.97	0.859	3.96	0.799	0.914	-

*p < .05 ** p < .01 *** p < .001

5 Discussion

Results from data analysis brought us to accept 3 hypotheses out of 5 hypotheses tested. For what concerns not accepted hypotheses, we cannot affirm that procedures that regulate physical access to IT resources (Hyp. 2), nor communications protocols (Hyp. 5), have a major effectiveness when outsourcing or managing in-house. The accepted hypotheses (shown in Fig. 2) confirm the major effectiveness of disciplinary processes when cybersecurity is outsourced (Hyp. 1), and the major effectiveness of protection of log information (Hyp. 3) together with the incorporation of lessons learned into recovery plans (Hyp. 4) when cybersecurity is managed in-house.

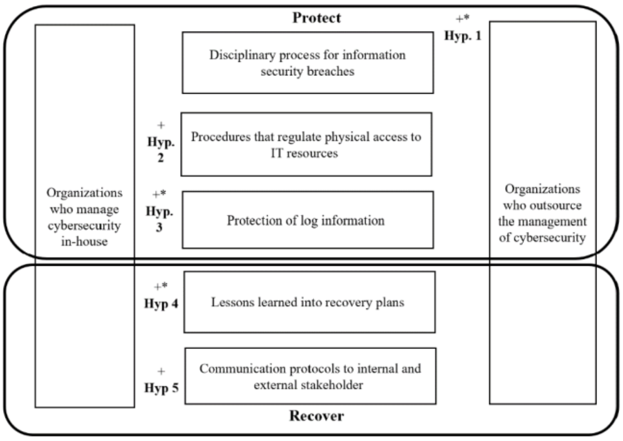

Protect
+*
Hyp. 1
Disciplinary process for information security breaches
+
Hyp. 2
Procedures that regulate physical access to IT resources
Organizations who manage cybersecurity in-house
+*
Hyp. 3
Protection of log information
Organizations who outsource the management of cybersecurity
+*
Hyp 4
Lessons learned into recovery plans
+
Hyp 5
Communication protocols to internal and external stakeholder
Recover

Fig. 2. Theoretical model and supported hypotheses

Even though disciplinary processes proved to be an effective practice for cybersecurity, what emerged from the results is that this practice should be adopted and/or encouraged particularly in those cases when the management of cybersecurity is outsourced. On the other hand, when organizations decide to manage cybersecurity internally, the relevance of the protection of log information and the inclusion of lessons learned into recovery plans confirms the importance for companies to focus more on building a sense of responsibility and awareness in employees, rather than put in action measures to discourage wrong behaviors. Companies should build cybersecurity by focusing on the awareness and sense of responsibility of their employees: the great majority of cyber threats nowadays aims at exploiting weaknesses and mistakes of human users, which confirms the important role that people, rather than advanced IT instruments, still have in ensuring cyber resilience for companies.

6 Conclusions and future work

This study has some limitations that suggest future research directions. First, the sample is limited to Italian cybersecurity experts; future research should increase the sample of respondents to include different national contexts and reduce the unbalanced nature of some organizational variables such as the one related to business to business/business to customer organizations. In addition, it would be interesting to expand the number of context-based variables, for example studying the effectiveness of NIST practices in public organizations. Finally, future research should also consider investigating other management practices to add to those already addressed by the NIST framework.

This research tested that disciplinary processes in case of non-compliance with company policies are more effective for organizations that decide to outsource the cybersecurity process, and the use of protection of log information and lessons learned is more effective for organizations that decide to manage internally the cybersecurity process. By confirming those differences, this study contributes to the cybersecurity management literature stream. The investigation of the context-based variables affecting cybersecurity effectiveness also provides managerial implications for managers belonging to organizations managing cybersecurity in-house or outsourced suggesting them what practices to encourage more based on their specific context.

Acknowledgement: This work is supported by the fund of the Department of Computer, Control and Management Engineering "Antonio Ruberti" of Sapienza University of Rome. The department has been designated by the Italian Ministry of Education (MIUR) for being “Department of Excellence” in advanced training programs in the field of cybersecurity.

References

[1] Z. A. Collier, D. Dimase, S. Walters, M. M. Tehranipoor, J. H. Lambert, and I. Linkov, “Cybersecurity standards: Managing risk and creating resilience,” *Computer*, vol. 47, no. 9. IEEE Computer Society, pp. 70–76, Sep. 01, 2014, doi: 10.1109/MC.2013.448.

[2] E. G. Carayannis, E. Grigoroudis, S. S. Rehman, and N. Samarakoon, “Ambidextrous Cybersecurity: The Seven Pillars (7Ps) of Cyber Resilience,” *IEEE Trans. Eng. Manag.*, vol. 68, no. 1, pp. 223–234, Feb. 2021, doi: 10.1109/TEM.2019.2909909.

[3] “Framework for Improving Critical Infrastructure Cybersecurity, Version 1.0,” Gaithersburg, MD, Feb. 2014. doi: 10.6028/NIST.CSWP.02122014.

[4] D. Di Mase, Z. A. Collier, K. Heffner, and I. Linkov, “Systems engineering framework for cyber physical security and resilience,” *Environ. Syst. Decis.*, vol. 35, no. 2, pp. 291–300, Jun. 2015, doi: 10.1007/s10669-015-9540-y.

[5] A. A. Ganin *et al.*, “Multicriteria Decision Framework for Cybersecurity Risk Assessment and Management,” *Risk Anal.*, vol. 40, no. 1, pp. 183–199, Jan. 2020, doi: 10.1111/risa.12891.

[6] H. Li, X. (Robert) Luo, J. Zhang, and R. Sarathy, “Self-control, organizational

context, and rational choice in Internet abuses at work," *Inf. Manag.*, vol. 55, no. 3, pp. 358–367, Apr. 2018, doi: 10.1016/j.im.2017.09.002.

[7] A. Annarelli, F. Nonino, and G. Palombi, "Understanding the management of cyber resilient systems," *Comput. Ind. Eng.*, vol. 149, Nov. 2020, doi: 10.1016/j.cie.2020.106829.

[8] A. Shah, R. Ganesan, S. Jajodia, and C. A. M. Hasan, "An outsourcing model for alert analysis in a cybersecurity operations center," *ACM Trans. Web*, vol. 14, no. 1, Jan. 2020, doi: 10.1145/3372498.

[9] J. Saleem, B. Adebisi, R. Ande, and M. Hammoudeh, "A state of the art survey - Impact of cyber attacks on SME's," in *ACM International Conference Proceeding Series*, Jul. 2017, vol. Part F130522, doi: 10.1145/3102304.3109812.

[10] "ENISA Threat Landscape Report 2018 — ENISA." https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2018 (accessed Jan. 29, 2021).

[11] S. Armenia, M. Angelini, F. Nonino, G. Palombi, and M. F. Schlitzer, "A dynamic simulation approach to support the evaluation of cyber risks and security investments in SMEs," *Decis. Support Syst.*, vol. 147, 2021, doi: 10.1016/j.dss.2021.113580.

[12] Panemon, "2018 State of Cybersecurity in Small & Medium Size Businesses," 2018.

[13] C. Ikerionwu, D. Edgar, and E. Gray, "The development of service provider's BPO-IT framework," *Bus. Process Manag. J.*, vol. 23, no. 5, pp. 897–917, 2017, doi: 10.1108/BPMJ-10-2015-0146.

[14] J. Doran, G. Ryan, J. Bourke, and F. Crowley, "In-house or outsourcing skills: How best to manage for innovation?," *Int. J. Innov. Manag.*, vol. 24, no. 1, Jan. 2020, doi: 10.1142/S1363919620500103.

[15] M. Benaroch, "Cybersecurity Risk in IT Outsourcing—Challenges and Emerging Realities," Springer, Cham, 2020, pp. 313–334.

[16] ISO/IEC 27001:2013, "ISO - ISO/IEC 27001:2013 - Information technology — Security techniques — Information security management systems — Requirements," 2013. https://www.iso.org/standard/54534.html (accessed Jan. 19, 2021).

[17] T. Herath and H. R. Rao, "Encouraging information security behaviors in organizations: Role of penalties, pressures and perceived effectiveness," *Decis. Support Syst.*, vol. 47, no. 2, pp. 154–165, May 2009, doi: 10.1016/j.dss.2009.02.005.

[18] J. D'Arcy, A. Hovav, and D. Galletta, "User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach," *Inf. Syst. Res.*, vol. 20, no. 1, pp. 79–98, 2009, doi: 10.1287/isre.1070.0160.

[19] L. Li, W. He, L. Xu, I. Ash, M. Anwar, and X. Yuan, "Investigating the impact of cybersecurity policy awareness on employees' cybersecurity behavior," *Int. J. Inf. Manage.*, vol. 45, pp. 13–24, Apr. 2019, doi: 10.1016/j.ijinfomgt.2018.10.017.

[20] B. Y. Ng, A. Kankanhalli, and Y. (Calvin) Xu, "Studying users' computer security behavior: A health belief perspective," *Decis. Support Syst.*, vol. 46, no. 4, pp. 815–825, Mar. 2009, doi: 10.1016/j.dss.2008.11.010.

[21] P. Ifinedo, "Understanding information systems security policy compliance: An integration of the theory of planned behavior and the protection motivation theory," in *Computers and Security*, Feb. 2012, vol. 31, no. 1, pp. 83–95, doi: 10.1016/j.cose.2011.10.007.

[22] M. Siponen, M. Adam Mahmood, and S. Pahnila, "Employees' adherence to information security policies: An exploratory field study," *Inf. Manag.*, vol. 51, no. 2, pp. 217–224, Mar. 2014, doi: 10.1016/j.im.2013.08.006.

[23] S. Vrhovec and A. Mihelič, "Redefining threat appraisals of organizational insiders and exploring the moderating role of fear in cyberattack protection motivation," *Comput. Secur.*, vol. 106, 2021, doi: 10.1016/j.cose.2021.102309.

[24] G. Dhillon, R. Syed, and F. de Sá-Soares, "Information security concerns in IT outsourcing: Identifying (in) congruence between clients and vendors," *Inf. Manag.*, vol. 54, no. 4, pp. 452–464, Jun. 2017, doi: 10.1016/j.im.2016.10.002.

[25] N. B. Akhuseyinoglu and J. Joshi, "A Risk-Aware Access Control Framework for Cyber-Physical Systems," in *Proceedings - 2017 IEEE 3rd International Conference on Collaboration and Internet Computing, CIC 2017*, Dec. 2017, vol. 2017-January, pp. 349–358, doi: 10.1109/CIC.2017.00052.

[26] Y. Cao, Z. Huang, Y. Yu, C. Ke, and Z. Wang, "A topology and risk-aware access control framework for cyber-physical space," *Front. Comput. Sci.*, vol. 14, no. 4, Aug. 2020, doi: 10.1007/s11704-019-8454-0.

[27] M. R. Doomun, "Multi-level information system security in outsourcing domain," *Bus. Process Manag. J.*, vol. 14, no. 6, pp. 849–857, 2008, doi: 10.1108/14637150810916026.

[28] A. Ahmad, S. B. Maynard, and G. Shanks, "A case analysis of information systems and security incident responses," *Int. J. Inf. Manage.*, vol. 35, no. 6, pp. 717–723, Dec. 2015, doi: 10.1016/j.ijinfomgt.2015.08.001.

[29] A. Cezar, H. Cavusoglu, and S. Raghunathan, "Outsourcing information security: Contracting issues and security implications," *Manage. Sci.*, vol. 60, no. 3, pp. 638–657, Sep. 2014, doi: 10.1287/mnsc.2013.1763.

[30] F. Fui-Hoon Nah, J. Lee-Shang Lau, and J. Kuang, "Critical factors for successful implementation of enterprise systems," *Business Process Management Journal*, vol. 7, no. 3. pp. 285–296, Aug. 01, 2001, doi: 10.1108/14637150110392782.

[31] L. K. Comfort and T. W. Haase, "Communication, Coherence, and Collective Action: The Impact of Hurricane Katrina on Communications Infrastructure," *Public Work. Manag. Policy*, vol. 10, no. 4, pp. 328–343, 2006, doi: 10.1177/1087724X06289052.

[32] A. Wieland and C. M. Wallenburg, "The influence of relational competencies on supply chain resilience: A relational view," *Int. J. Phys. Distrib. Logist. Manag.*, vol. 43, no. 4, pp. 300–320, May 2013, doi: 10.1108/IJPDLM-08-2012-0243.

[33] K. Scholten and S. Schilder, "The role of collaboration in supply chain

resilience," *Supply Chain Manag.*, vol. 20, no. 4, pp. 471–484, Jun. 2015, doi: 10.1108/SCM-11-2014-0386.

[34] R. Knight and J. R. C. Nurse, "A framework for effective corporate communication after cyber security incidents," *Comput. Secur.*, vol. 99, p. 102036, Dec. 2020, doi: 10.1016/j.cose.2020.102036.

[35] "Decreto direttoriale 6 novembre 2019 - Elenco dei manager qualificati e delle società di consulenza," 2019. https://www.mise.gov.it/index.php/it/normativa/decreti-direttoriali/2040421-decreto-direttoriale-6-novembre-2019-elenco-dei-manager-qualificati-e-delle-societa-di-consulenza (accessed Jan. 29, 2021).

[36] J. R. Warmbrod, "Reporting and Interpreting Scores Derived from Likert-type Scales," *J. Agric. Educ.*, vol. 55, no. 5, pp. 30–47, Dec. 2014, doi: 10.5032/jae.2014.05030.

[37] D. Cramer, *Fundamental statistics for social research: Step-by-step calculations and computer techniques using SPSS for Windows.* Taylor and Francis, 2003.

[38] A. P. Aldya, S. Sutikno, and Y. Rosmansyah, "Measuring effectiveness of control of information security management system based on SNI ISO/IEC 27004: 2013 standard," in *IOP Conference Series: Materials Science and Engineering*, Aug. 2019, vol. 550, no. 1, doi: 10.1088/1757-899X/550/1/012020.

[39] J. Gutiérrez-Martínez, M. A. Núñez-Gaona, and H. Aguirre-Meneses, "Business Model for the Security of a Large-Scale PACS, Compliance with ISO/27002:2013 Standard," *J. Digit. Imaging*, vol. 28, no. 4, pp. 481–491, Aug. 2015, doi: 10.1007/s10278-014-9746-4.

6. Leveraging human-machine interaction for cyber resilience

This chapter will address RQ4.1. and RQ4.2. by exploring how cyber resilience can benefit from increased and fruitful human-machine interaction. Specifically, the thesis focused on the topic of digital assistants as a tool to improve performance while supporting the operator in alienating or complex operations. Section 6.1 presents an under-review article that develops a conceptual architecture and taxonomy to guide researchers in the development of digital assistants. The work went to fill the void generated by a lack of common consensus on which should be technical and functional characteristics of a virtual assistant. Section 6.2 reports on the development of a digital assistant for cybersecurity. In addition to recounting the development of the agent's functionality, the research will focus on analyzing the benefits and limitations of such a solution, trying to answer the increasingly ongoing question: can digital assistants, and in a broader view human-computer interaction, help close the gaps in cyber resilience?

As mentioned previously, these concepts fit into the new paradigm for the industrial transformation described by Industry 5.0. Industry 4.0 was a technologically focused and growth-oriented industrial paradigm that failed to consider the environmental, sociological, and long-term development components of economic operations [251], [252]. Similarly, building resilience within our present economy and adapting it to be more resilient to future shocks, suggests that mitigating the consequences of disruptions (e.g., a pandemic or a data breach) should be more ambitious than trying to return to baseline conditions. New systemic solutions are required. Industry 5.0 tries to provide a solution by stressing the importance of: adopting a human-centric approach to digital technologies including artificial intelligence; up-skilling and re-skilling workers on digital skills; sustaining modern, resource-efficient, and sustainable industries and transitioning to a circular economy.

Among the technologies used as an interface for human-machine communication are digital assistants. Also known under other terms such as chatbot, voice-enabled assistant, intelligent assistant, and conversational agent, this technology is among the fastest-growing information technology applications [253]. This solution has been used in various fields from marketing to healthcare or the industrial sector with different applications such as support in maintenance procedures, assembly, or quality control. The inclusion of such applications has brought benefits such as supporting the operator in safety-critical, complex, and high-precision operations [254], [255] but also relieving the operator from stressful, repetitive, and alienating operations.

On the other hand, such an application requires cautious onboarding that includes, for instance: Actively engaging employees in business process design, effective employee training, ai-focused change management, and support to reduce risks such as cybersickness or employees' natural fear that automation will take their jobs away [256], [257].

For more details on these aspects and a prototype solution, please refer to the following two articles prepared during the PhD pathway and currently under review. The Appended Paper 3 is currently under review at the Journal of Industrial Information Integration. The Appended Paper 4 is currently under review at Expert Systems With Applications.

6.1. Appended Paper 3 Under Review: Human-technology integration with industrial conversational agents: a conceptual architecture and a taxonomy for manufacturing.

Human-technology integration with industrial conversational agents: a conceptual architecture and a taxonomy for manufacturing

Highlights

- We designed a conceptual architecture and taxonomy for developing industrial conversational agent
- We fill the need of empirical research on conversational agents in manufacturing
- We present a qualitative analysis of conversational agents in manufacturing
- We highlight the importance of positive and beneficial human-machine interaction.

Abstract

Conversational agents are systems with great potential to enhance the human-computer interaction in industrial settings. Although the number of applications of conversational agents in many fields is growing, there is no shared view of the elements to design and to implement chatbots in the industrial field. The paper presents the combination of many research contributions into an integrated conceptual architecture, for developing industrial conversational agent using the Nickerson's methodology. The conceptual architecture consists of five core modules; every module consists of specific elements and approaches. Furthermore, the paper defines a taxonomy from the study of empirical applications of manufacturing conversational agents. Indeed, some applications of chatbots in manufacturing are available but those have never been collected in single research. The paper fills this gap analyzing the empirical cases and presenting a qualitative analysis, with verification of the proposed taxonomy. The contribution of the article is mainly to illustrate the elements needed for the development of a conversational agent in manufacturing: researchers and practitioners can use the proposed conceptual architecture and taxonomy to more easily investigate, define, and develop all the elements for chatbot implementation.

Keywords: chatbot; natural language processing; dialogue systems; voice bot

1. Introduction

Conversational agents belong to the systems designed to enable Human-Computer Interaction [1]. These systems represent a new form of interaction between humans and machines, allowing the user to interact using the tool most used by humans: natural language [2]. These interfaces represent a paradigm shift from the current Graphical User Interfaces (GUIs), where interaction is based on a visual representation that includes elements such as icons, sliders, and buttons [3]. The objective of these new interfaces is to offer a new, logical and intuitive human-computer interaction by representing a cost-effective solution that can facilitate, speed up and increase the efficiency of daily activities [4]. This allows users to intuitively interact with data, resources and services without the need for GUI training: the user can simply make a request through the use of their own language, and be assisted and supported by the conversational agent [5].

With the term conversational agents are indicated all those software able to support a conversation with a human being through a textual and/or vocal channel. In literature are used multiple terms to indicate such systems, including: conversational systems, conversational user interfaces, chatbot, voice assistant, virtual assistant, spoken dialogue system, conversational AI [6]. Although there are some differences, the term chatbot is by far the most used to refer to such solutions, terminology that should be intended in its most general definition of conversational agent [5]. Thus, in the paper, the authors use chatbot and conversational agent as synonyms. In the manufacturing sector, the adoption of conversational agents is driving the digital transformation of organizations, with the aim of improving both customer and user-experience and making their internal processes more efficient [7]. These technologies are included in the broader scope of eXtended Reality (XR) technologies, which are leading the way towards new forms of interaction with computers. Their goal is to increase the degree of mobility, autonomy and independence of operators by working on Human-In-The-Loop, user-centered systems, in which operators play the role of decision makers, entrusting the most repetitive operations to these technologies [8]. With this in mind, the development of conversational agents is focused on both supporting users in interacting with machines [9], databases [10], information systems [11], and in completing tasks [12], moving towards the notion of smart operators [13]. It is to underline that conversational agents require a proper design even to cope with possible safety issues, that are always present in 4.0 technologies [14,15], because the increasing introduction of digitalization and automation of work processes lead to the expanded complexity of cyber-socio-technical systems[16].

From the analysis of the few papers devoted to conversational agents in the industrial field, there is no agreement on the elements to be considered and developed for the creation of an industrial conversational agent. In this paper, we evaluate the key elements specific for the industrial conversational agents and we review the literature to build an integrated architecture for developing industrial chatbots.

Therefore, this paper addresses the following research questions:

- RQ1: Which logical interconnections and modules are needed for a conversational agent’s architecture?
- RQ2: What are conceptually grounded and empirically validated design elements for manufacturing conversational agents?

To answer RQ1 architectures available in literature are first investigated. Then, the research presents and discusses the fundamental concepts for understanding the logical operations of an industrial conversational agent through the definition of its general design and its modules, to propose an architecture, we assume “integrated” since it integrates several literature contributions. Subsequently, the attention has been turned towards the development and use of such systems in the manufacturing sector, analyzing their role as an enabling technology for Industry 4.0/5.0. For this purpose, a reference taxonomy was developed to answer RQ2. The research approach for its development follows a revised and adapted version of the taxonomy development model of [17]. The taxonomy is then used to classify a sample of 20 manufacturing chatbots, appropriately selected from various scientific databases. The classification confirmed the validity of the taxonomy and underlined main paths in up-to-date manufacturing conversational agents.

The paper is organized as follows. Section 2 introduces the topic of conversational agents providing literature background information. More specifically chatbot architectures and technical terminology available in literature are underlined. Section 3 describes the conversational agent conceptual architecture for industry. Section 4 details the research process followed to develop the taxonomy and presents it. Section 5 provides an extensive case study qualitative analysis using the proposed manufacturing chatbot taxonomy. Finally, section 6 concludes and outlines the follow-up research.

2. Related work and motivation

Although the interest for conversation systems has increased in recent years both in industry and in research [7], the idea of applications capable of interacting with humans was born in 1950, when Alan Turing wondered if machines were able to "think", to link and express ideas [18]. In 1966, Joseph Weizenbaum [19] created ELIZA, which has been historically considered as the first conversational system. A first generation of conversational agents whose operation was based on the use of specific rules was developed starting from ELIZA. PARRY (1972) is considered the first chatbot with personality and ALICE (1995) the first chatbot to be developed with the Artificial Intelligence Mark-Up Language (AIML) [20]. Such systems have seen a significant evolution in recent years due to advances made in the field of Artificial Intelligence (AI). On one side, Natural Language Processing (NLP) techniques have allowed for better syntactic and semantic analysis of text [21] with application in several fields [22]. On the other side, Machine Learning applications have allowed for a move away from rule-based implementation, leading systems to learn directly from large corpus of data [20]. The explosion of such technologies then occurred with Apple's introduction of Siri in 2010 and followed by Watson Assistant, Alexa, Cortana, and Google Assistant [23].

This widespread use has led to the theorization of multiple reference architectures and functionalities for the development of conversational agents. The logical functioning of a generic conversational agent can be schematized as follows: once it receives the user's input, the system analyses it using Natural Language Processing techniques to identify what the user wants to obtain. Once the chatbot has identified the correct Intent, it must provide the correct or best possible answer by performing one of the corresponding actions [24].

Among the most straightforward architectures is the one proposed by McTear (2020). Despite it is not highly detailed, this architecture applies well to both text-based and voice-based chatbots. The main difference is that the latter type will be equipped with a speech recognition module to process the voice input provided by the user and a text-to-speech module to transform the chatbot output into voice format. Among other researches that provide a complete chatbot design architecture is the one by Adamopoulou & Moussiades [23] and more recently the one by Serras et al. (2020) that integrates this work also with extended reality (XR) components. Overall, five fundamental modules return across these designs: Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), Dialog Manager (DM), Natural Language Generation (NLG), and Text-To-Speech (TTS).

In terms of functionality, chatbots mainly fall into two different categories : Task-Oriented and Non-task oriented chatbots [5]. In Task-Oriented, the interaction between humans and machines is focused on accomplishing a specific task. They are designed to deal with a specific scenario and perform best with a narrow knowledge domain. On the other hand, non-task oriented are designed to have more extended conversations, with the goal of simulating a real conversation between humans. They often have recreational, or entertainment purposes and the conversations are based on a broader knowledge domain. A few authors further subdivide this category into Informative and Conversational. The former are intended to provide the user with specific information (FAQbot, Q&Abot), the latter are intended to hold generic conversations with users [7]. Further classifications in the literature concern the method of response generation, the knowledge domain, the length of the conversation, the service provided, and the control of the conversation. A distinction is made between the Rule-based Approach and the Neural Network Based Approach, which in turn is divided into retrieval based approach and generative approach [26]. Sometimes in the literature the terms Rule-based chatbot and Data-driven chatbot (or AI-based chatbot) are also used to indicate the different types of chatbot that can be realized [5]. Classification by knowledge domain is related to the amount of available data, which constitutes the chatbot's knowledge base. One can distinguish Open domain and Closed domain chatbots [20]. When talking about Open domain, the conversation with the chatbot can start in one knowledge domain and later move to a different one. In contrast, Closed domains have limited knowledge about a specific domain and are designed to have conversations focused on one or a few specific topics [27]. Based on the length of the conversation, two other types of chatbots can be distinguished: systems based on Short-Term and Long-term relations [1]. A short-term relation is characterized by a one-shot interaction, also called single-turn [28], in which the response is generated solely based on a single message, without collecting the user's information. In contrast, Long-term, also called multi-turns, are chatbots designed to have an extended interaction over time and able to record relevant information exchanged during the conversation. Furthermore, in user-chatbot interactions, two categories are distinguished based on who drives the dialogue: chatbot-driven dialogue and user-driven dialogue systems [4]. Finally, conversational agents can be classified according to the type of relationship with the user and the type of service they provide [7,23]. Interpersonal chatbots have the sole purpose of giving the requested information and moving on to the next user. Intrapersonal, on the other hand, are those chatbots that have an elevated level of engagement with the users, also performing tasks for them.

As discussed, there are several criteria for classifying chatbots in the literature. These classification criteria should not be understood as mutually exclusive. Two or more criteria may coexist and be used in. combination for the development of a chatbot. Although this is typically the scenario, there are logical relationships between these criteria that must be considered. When designing a chatbot, the options to be implemented depend on its ultimate purpose. Based on the final aim of the chatbot there will be advisable, viable, and avoidable options considering functional suitability, performance efficiency, usability, and security [7]. However, despite diverse chatbot characteristics that have been investigated, there is a scarcity of empirical research on how to design chatbots’ profile. Notably as reported in the survey by Motger et al. [7] there is a lack of structured and synthesized knowledge. They underlined as one of the major challenges in the field of conversational agents is the shift from develop chatbots for simple tasks moving towards assistants able to perform complex tasks by applying domain and target specific requirements. This is particularly relevant in the manufacturing sector where the topic of conversational agents is still in its beginning phase, cases presented are unstructured, lacking a common line for their development, evolution, and personalization.

Therefore, from that reviewed, the objective of the current work is to determine which are conceptually all the design elements for a manufacturing chatbot and to address a taxonomy and guideline for its development. The taxonomy will be based on scientific literature and validated through empirical data collected from real manufacturing chatbot case studies.

3. Conversational Agent Conceptual Architecture

An appropriate conversational agent architectural design is the first step to investigate for the development of a chatbot. Therefore, several architectural designs have been proposed in literature. Some of them have been approach specific. For instance in [29] and [30] the authors propose architecture specific for rule based chatbots and retrieval based chatbots. In their review illustrate specific architectures for corpus based, intent based, or recurrent neural network based chatbot. Other have been function specific such as [31] which has focused on architecture modules for human-computer speech interaction.

Among the first design is the one by Souvignier et al. [32] who present a system architecture focusing on elements fundamental for spoken dialog systems. Their main components are a speech recognizer, a natural language understanding module, a text-to-speech tool, and a dialog manager. The research offers a detailed technical description of the natural language understanding module but lacks in other architectural details and there is no technical information on Speech recognition, Speech Synthesis and Response Generation.

Among the most extensive and complete recent chatbot architecture is the one proposed by Adamopoulou & Moussiades [23]. Their work offers both an architecture and a development approach. Despite its interesting integration of different modules, their design lacks in details regarding Natural Language Understanding techniques, Dialog Policies categories and the Response Generation Component lack many essential details. An interesting design is proposed by Serras et al. [25] who propose an Interactive XR architecture structured in layer. It integrates a spoken dialogue module along with a Device Control Layer, an Interpretation Layer, Domain Knowledge Layer and Response Generation Layer. However, it is quite abstract as it does not provide essential details for each layer, especially the dialogue manger module has not been articulated in its submodules. Besides [28] and [5] present two simplistic designs that on one side lack of many essential details but on the other side offer two clear and straightforward approaches for the definition of the main modules a chatbot must have. All the main components are then described in detailed focusing on task oriented and rule-based dialogue systems development. Finally, this review of chatbot architecture literature has demonstrated an absence of terminological consistency. Terms such as Natural Language Understanding [5,28], Spoken Language Understanding and Semantic Codification [25] or User Message Analysis [20] are used as synonymous. Instead, the term Dialog Manager is widely used, with some differences such as Dialogue State Tracking [28] or Dialogue Policy

Optimization [24]. Similarly, Natural Language Generation [5,28], Response Generation Component or Layer [20,25] are used.

In this section, the authors compose a shared architecture that considers those developed to date, offering an articulated path between the various architectural steps, with a detail on each phase and a terminological consistency. The architectural design is at the same time general and detailed including all the modules from the beginning of the conversation to the response generation. The proposed architecture is shown in Figure 1. It is characterized by 5 core modules, explained below: Automatic Speech Recognition (ASR), Natural Language Understanding (NLU), Dialog Manager (DM), Natural Language Generation (NLG), and Text-To-Speech (TTS).

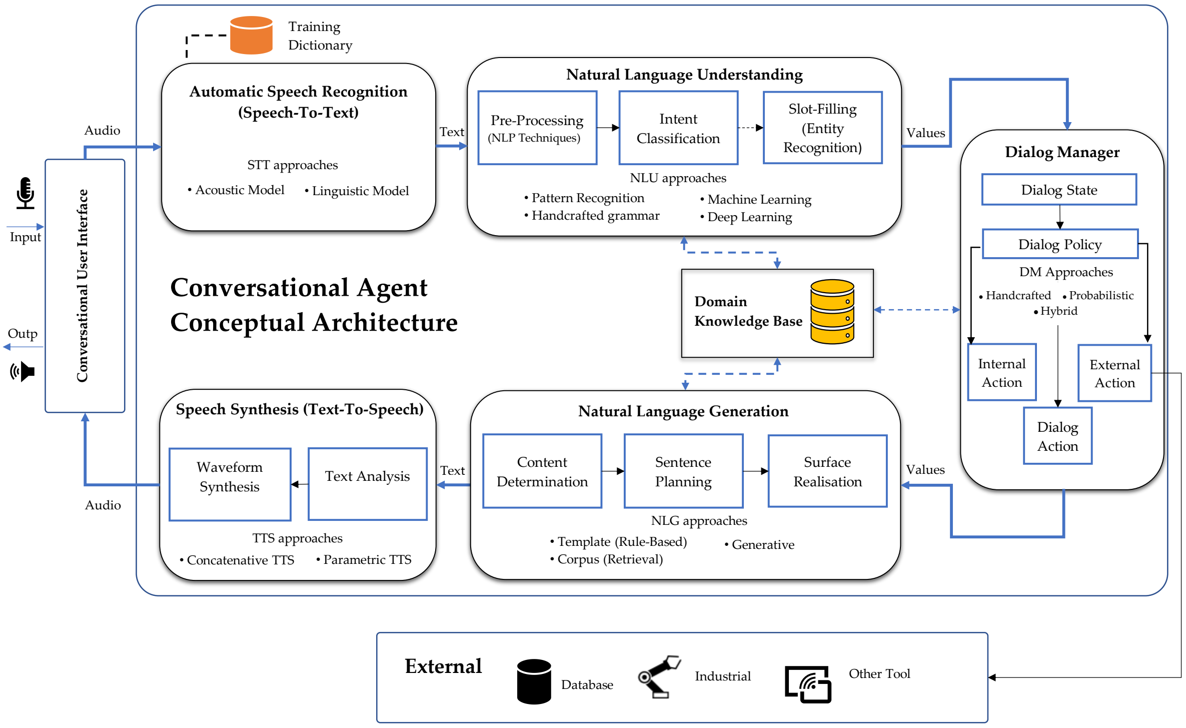


Figure 1 - Conversational Agent Conceptual Architecture

Automatic Speech Recognition

The first module is the Speech-to-text. Its task is to capture and transcribe in text format the vocal input given by the user. The purpose is to collect a set of data to be processed by the NLU. Modern ASRs are based on the combination of two probabilistic models: the acoustic model, which calculates the most probable sequence of phonemes corresponding to each part of the speech signal; and the linguistic model, which calculates the most probable sequence of words that match the previously calculated sequence of phonemes [33]. The main goal is to minimize the Word Error Rate. The most used techniques are based on Deep Neural Networks, such as Long-Short-Term-Memory [34] and Hidden Markov Models [35], which allow to achieve a word error rate below 10% [33].

Natural Language Understanding

The Natural Language Understanding module is responsible for analyzing the string provided by the ASR to determine its meaning [5]. It is the process of transforming sentences into structured information. Specifically, two basic functions can be performed in the NLU module: Intent Classification (or Intent detection) and Slot Filling (or entity recognition). The process of text comprehension begins with the use of NLP techniques. The main ones are Token decomposition (Tokenization); morphological and lexical analysis through Part-Of-Speech (POS); syntactic analysis through the generation of a Parse Tree. Other techniques that can be used are Lemmatization, Stemming, and Sentiment Analysis. Once the text has passed the NLP phase, it proceeds with intent classification and eventually slot-filling. These functions can be performed following rule-based approaches or machine learning. Early chatbots were based on pattern matching algorithms [20]. These involve the creation of several categories, each with corresponding patterns and templates. The user's phrases are then matched with a pattern and the content of the template is given in response. The major issue with this approach is the required perfect match between input and pattern. Another type are rule-based chatbots. These are used to extract context, intent, and slots from the user's sentence in order to match certain keywords, using Handcrafted Grammars [5]. HGs contain all the rules required to cover the expected user inputs, adding a degree of flexibility of possible inputs over pattern matching. They also involve specific rules for each input by requiring different rules for sentences having the same meaning but a different structure. To date, the most widely used technique for NLU is the use of Machine Learning methods to extract intents and slots from user inputs. With this approach, the NLU module requires a corpus, i.e., a set of sentences, used to train the chatbot. For each intent, a list of training utterances is provided, on which the chatbot is trained. In this approach, the identification of a phrase in a specific intent is treated as a classification problem and supervised Machine Learning algorithms are used. This approach is more robust than Handcrafted grammars; in fact, inputs can be linked to an intent even when the sentence wording is not the same as the examples in the corpus. Moreover, conversational agent using machine learning techniques are also characterized by slot filling capabilities. With slot filling the system continuously parses the user's responses for information that it uses to guide the conversation. This means the agent can recognize information that the user has already provided or that is missing, ask clarifying questions if needed, and continue with the dialog. Finally, recently the use of Deep Learning and neural networks (Recurrent Neural Network) has become more widespread, mainly employed for the development of generative chatbots [24,26].

Dialog Manager

The DM is the core module of a conversational agent, it manages the conversation and decides, at each iteration, which actions must be performed based on the input (Intent) provided by the user. It manages the conversation with the user to achieve the goal expressed. The module consists of two main components [36]: Dialog State and Dialog Policy. The Dialog State tracks Intent and slots and is updated at each user iteration. The Dialog Policy is the strategy aimed at acquiring the missing slots to correctly complete the query [33]. Here the system decides the action to be taken based on what is reported in the dialog state. Depending on the moment of the conversation, 3 different types of actions can be performed in the dialog policy: dialog, external and internal action. Dialog actions correspond to a message sent to the user in response to his request and allow the dialogue with the user. They can be a confirmation action, a request for further information or an answer to the user's query. External actions are actions that allow the conversational agent to interact with services provided by other software or databases to satisfy the user's request (e.g., activate robots or extract information). Finally, Internal Actions are actions that the agent uses to modify its behavior and improve its performance. Ultimately, the approaches used for the development of DM, and in particular Dialog Policy, fall mainly into 3 categories: handcrafted, probabilistic, hybrid [36] depending on the possible states and transitions between states of the conversation. The Handcrafted Approach defines both the state of the system and its policy through a set of rules that establish the state of the conversation and which actions are possible for each state. In the Probabilistic Approach, the system learns the rules from real conversations (from a corpus). The corpus contains examples of responses and conversations. Specifically, corpus-based chatbots select the most correct answer by matching the user's request with an example contained in the corpus that is used as the answer. Finally, the Hybrid Approach combines the advantages of purely rule-based and data-driven approaches.

Natural Language Generation

The NLG module is responsible for generating the response text, based on the decision made by the DM. The DM communicates the relevant information contained in the dialog state to the NLG, which is responsible for structuring that information into words and sentences. The NLG module involves three processes: content determination, sentence planning, surface realization. Content determination is the process of deciding what information should be realized. This step has to deal with selection, abstraction and filtering of the input data removing irrelevant information. Sentence planning is the process of ordering and grouping the semantic information into chunks that are coherent and desirable. Finally surface realization is the process of placing the structure, relevant words and producing a well formed sentence that fits the rules of grammar. The most appropriate response is generated based on three different possible approaches: Rule-based, Retrieval, and Generative approach. In Rule-based, the response has a predefined structure and is contained in a specific template. Conversely, in Retrieval, the best possible answer is selected from a predefined corpus containing answer examples by Machine Learning algorithms. Finally, in Generative, the answer is completely generated by the chatbot by Deep Learning algorithms, not making use of any kind of predefined answers.

Text-To-Speech

The Text-to-speech or Speech Synthesis module is the last module that makes up the architecture of a conversational agent and is tasked with converting text generated by the NLG and synthesizing it to generate output in speech format [37]. In order to accomplish its task, the TTS module relies on two steps: Text Analysis, in which the text to be read is transformed into a representation consisting of phonemes and prosodic information, and Waveform Synthesis, in which the internal representation is converted into a waveform that can then be output as a voice message [5]. There are two specific methods for conversion: concatenative TTS and parametric TTS. In concatenative TTS appropriate "speech units" contained in a speech corpus are selected and concatenated to obtain the final waveform.

Parametric TTS instead uses digital signal processing technologies to synthesize speech from text. There are mainly two models used for concatenative TTS: one based on Linear Prediction Coefficients (LPCs) and the other based on Pitch Synchronous OverLap Add (PSOLA). As for parametric TTS, most used methods are Hidden Markov Models (HMMs) and Deep Neural Networks (DNNs) [37].

Based on the above, as can be seen from Figure 1, the architecture is grounded on a principle of close collaboration between modules which, while being independent, affect the performance of subsequent modules and operate in synergy. For example, training the Automatic Speech Recognition (ASR) module through an appropriate training dictionary, allows the NLU to simplify the process of identifying the intent and slots. On the other hand, a highly effective NLU module can make the DM perform better by shortening the duration of the conversation with the agents [36].

4. Taxonomy of design elements for manufacturing chatbots

4.1. Taxonomy development procedure

Conversational agents represent one of the solutions to drive organizations' digitization process. This technology offers potential support for various processes and activities within industrial plants with the aim of enabling a new degree of interaction, control, and efficiency. To date, there is a small number of applications, in literature a few application cases can be found ranging from operator assistance in production, maintenance, training and information collection. However, as far as the authors know, there are no specific taxonomies to support the selection of manufacturing chatbot’s elements.

In the literature two relevant taxonomies are referenced: the one by Janssen et al. [38] and the one by Nißen et al. [39] . However, their proposals focus on a taxonomy for closed-domain conversational agents with no reference to a particular domain and/or context. The manufacturing field, on the other hand, has specific characteristics, based on a task-oriented logic. These types of chatbots are designed with the goal of achieving a specific purpose and to assist the user in one or few specific tasks [7]. Such systems are short-conversation agents [40] and work through the execution of preconfigured actions oriented towards the achievement of a specific goal [41] in a closed domain with limited knowledge.

Starting from the most generic reference taxonomies and detailing them by exploiting application cases of chatbots in manufacturing we present below a taxonomy of design elements for manufacturing chatbots.

Its development has been done readapting the steps suggested by Nickerson's model [17]. It is based on the identification of meta-characteristics, i.e., those more general design dimensions that will be the basis for the choice of the final characteristics. Next, the model states that each dimension must be representative of a unique design element and be decomposed into at least two characteristics. These final characteristics will have to be mutually exclusive to classify a chatbot assigning only one characteristic for each dimension.

To answer our RQ2 we have adapted to the manufacturing scenario the steps proposed by Nickerson et al. [17]. The approach proposed combines: i) two chatbot taxonomies [38,39] previously developed following Nickerson approach; ii) task oriented conversational agent literature; iii) case studied and empirical data related to manufacturing chatbots.

Our taxonomy builds on existing general chatbot taxonomies developed respectively through six iterations [39] and seven iterations and 103 chatbot articles [38]. In iteration one taxonomies are merged and duplicates removed. In iteration two dimensions are revised adapting them to the manufacturing scenario which is conceptually focused on task oriented conversational agents. Iteration three revised the taxonomy based on empirical observations and case studies described in twenty manufacturing chatbot articles. Finally, through a conceptual to empirical third iteration the terms are revised according to the latest knowledge discussed in the domain related scientific literature.

Before starting describing each iteration, [17] recommends the definition of (i) a purpose of the taxonomy and the determination of (ii) meta-characteristics and (iii) ending conditions.

First, the purpose of our taxonomy is to provide a design taxonomy to guide researchers and practitioners in the development and comprehension of manufacturing conversational agents. Second, meta characteristic are defined. [17] defines them as the basis for the choice of taxonomy characteristics and underlines the importance of considering expected end users of the taxonomy. [39] focuses on the importance of human-like interactions proposing three related perspectives: intelligence, interaction and context. [38] instead stress on visible or experiential in human-chatbot interaction. Our scenario takes up the rationale of defining meta-features based on the concepts of machine interaction, however, believing that it is important in a production context to also provide the developer with a more technical perspective and not just interaction related. For this reason, we identified two perspectives: Chatbot perspective and Chatbot-User interaction perspective. The first one identifies all those design elements that directly concern the development of the chatbot and its functionalities. The second one refers to dimensions and features that qualify the interaction between chatbot and user. Regarding the selection of ending condition, this study adopted all objective and subjective conditions suggested by Nickerson et al. [17].

Iteration 1 – Conceptual to empirical: Merging chatbot taxonomies

The difference with [17] proposed approach can be traced to this iteration. Our research in fact restarts from the latest iterations of [38,39] works and from these restarts by customizing and extending their taxonomies. The study of the literature has shown how well these taxonomies describe the characteristics of conversational agents however when focusing on a specific domain these are not sufficient to guide the development of chatbots. In particular, the manufacturing context is characterized not only by strong human-chatbot interaction but also by a need for human-chatbot-machine coordination to be taken into account when developing chatbot conversations [13]. In addition, the objectives of chatbots in manufacturing are varied: training, operator assistance, data collection etc., and each of them needs a detailed definition of dimensions and characteristics.

Specifically, in this first iteration we reviewed [39] and [38] taxonomies and merged them to derive an initial set of design dimensions. Duplicates have been removed.

Iteration 2 – Conceptual to empirical: Refinement of the taxonomy for a task oriented prospective

As mentioned in the previous paragraphs, the analysis of the literature has underlined that in the manufacturing environment, task-oriented conversational agents find major application.

This class of systems are designed with the goal of achieving a well-defined purpose, and to assist the user in one or a few specific tasks [7]. In this second conceptual to empirical iteration, we aimed at analyzing each dimension and to evaluate to which extent they might be design-relevant for task-oriented chatbots.

Therefore, dimensions such as Application Domain, Collaboration Goal, Motivation for chatbot use, and Primary Communication Style, which are suitable in the reference taxonomies to identify the application domain and functionality of chatbots, are removed because they are representative of generic characteristics of Closed-Domain chatbots and do not meet the ultimate purpose of our taxonomy.

Iteration 3 – Empirical to conceptual: classification of manufacturing conversational agents' dimensions

For the third iteration, we chose an empirical-to-conceptual approach to customize the taxonomy on a manufacturing perspective. We have selected twenty published manufacturing chatbot case studies retrieved from three main scientific databases: Scopus, ResearchGate and Google Scholar. To determine the case studies, the search was done by keywords and then by analyzing articles cited in text and contributions that cited the selected cases. Each case study has been analyzed to identify which design dimensions and characteristics researchers focused on when developing a manufacturing chatbot. Each dimension and related characteristics identified was compared with existing ones to assess their similarity. Similar dimensions have been merged. Some characteristics have been revised or added. When no similar dimension was identified it was added as a new taxonomy dimension.

Figure 2 shows in detail all the dimensions added and the following paragraph will explain their meanings.

Iteration 4: Empirical to conceptual: Refinement of the taxonomy

In this iteration, we chose the empirical-to-conceptual path again. Dimensions names have been revised for a more complete understanding and to be consistent with manufacturing terminology. It was finally decided to leave some dimensions even though these were not found in the manufacturing articles. The choice was made by observing how in similar contexts in terms of human-machine interaction and process complexity (e.g., healthcare, cybersecurity) these dimensions have been used. Therefore, as explained in detail it was deemed important to leave these dimensions in the taxonomy. In this iteration no new dimensions have been added and all the ending conditions were fulfilled, and the taxonomy process was completed.

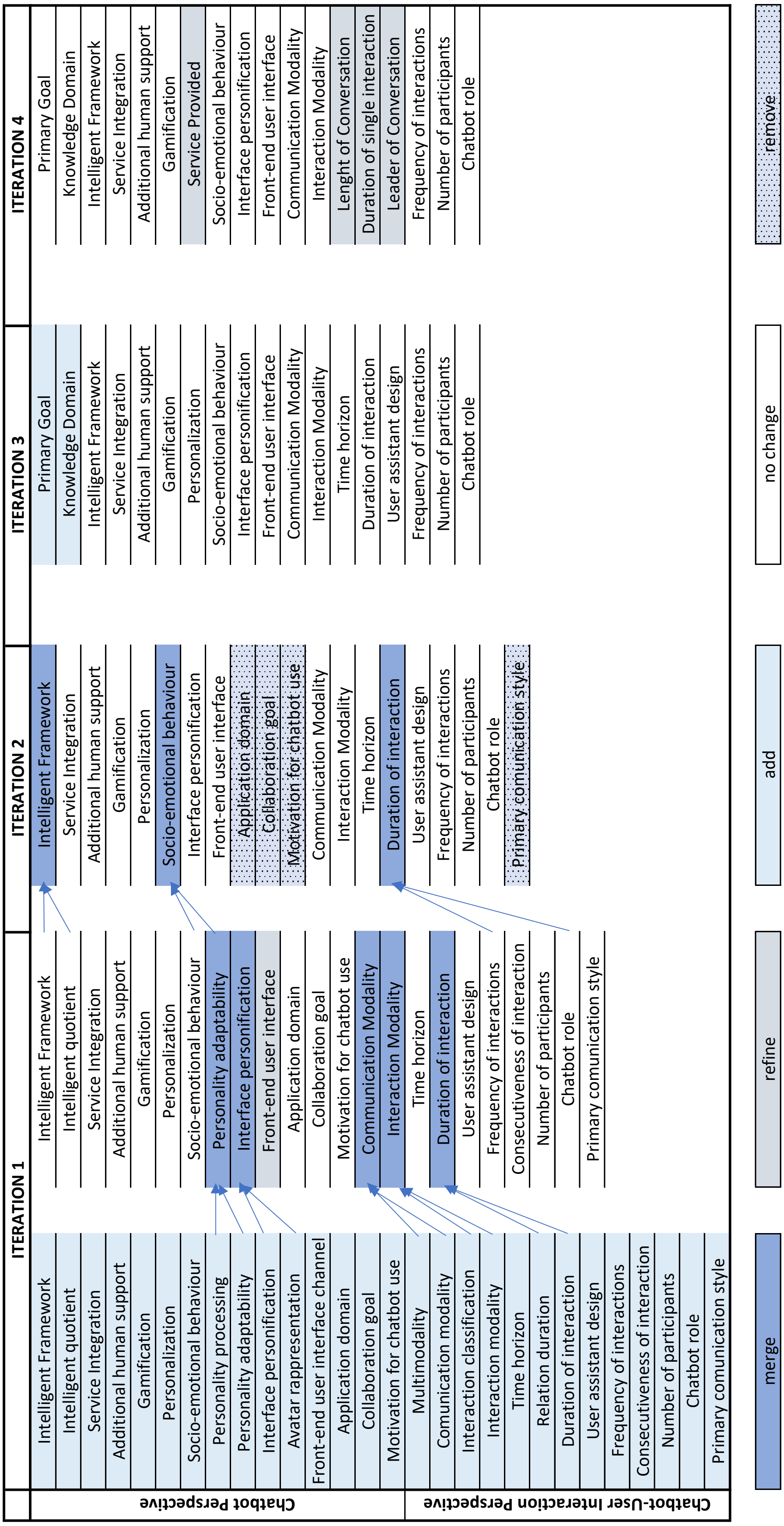


Figure 2 - Taxonomy development process

All the iterations which encompass the integration of reference taxonomies, conversational agents' literature, and the cross-reading of manufacturing chatbot application cases allowed to define the 18 design dimensions that make up the final taxonomy. Moreover, 42 characteristics have been determined, which can be divided into chatbot and chatbot-user interaction perspective. Table 1 shows the proposed taxonomy for task-oriented conversational agents in manufacturing. The following paragraph details each taxonomy dimensions.

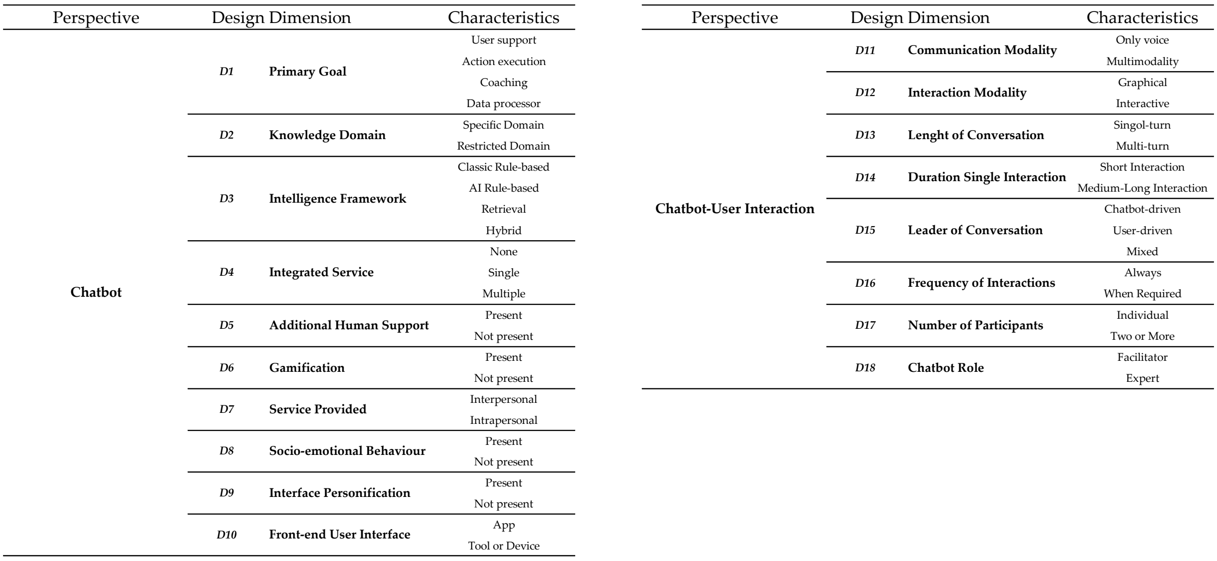
Chatbot perspective

The first dimension defined is D1 Primary Goal which defines the purpose of the chatbot. For task-oriented manufacturing chatbots, 4 primary goal characteristics can be distinguished in relation to the primary purpose for which the chatbot is implemented [7]: user support, action execution, data processor and coaching. User Support supports the user in the operational execution of their activities guiding them step by step with the aim of improving the user-experience. Data processors (or Information request) support the decision making process of operators by offering quick and easy access to corporate databases and collecting data for and from users [42]. Action execution enables control through voice commands of other integrated systems and software [43]. Finally, Coaching (or User training) are chatbots focused on training, evaluation and dissemination of corporate know-how [44].

D2 Knowledge Domain dimension refers to the extent of chatbot's knowledge domain. Through this dimension, the degree of specialization of conversational agents in manufacturing is analyzed by assessing how many different tasks or contexts it can handle within its closed domain. Depending on the extent of the knowledge domain, two categories are defined: Specific domain and Restricted domain. The first refers to chatbots with only one context or task defining its domain, such as LARRI [42] and Max [43]. The second instead refers to systems that can handle several different activities from each other, such as Bot-X [45] and chatbot coaching [44].

D3 Intelligence Framework dimension indicates the type of chatbot. A chatbot may be classified as Classic Rule-based, AI Rule-Based, and Retrieval. To date, these are the most widely used approaches for implementing chatbots in manufacturing. Specifically, this subdivision gives insight into the technical principles of chatbot development to understand and analyze user input (NLU), process information (DM), and select response (NLG). Figure 1 shows the differences of NLU, DM, and NLG according to the selected feature.

Table 1 - Taxonomy of design elements for manufacturing chatbots



Perspective	Design Dimension		Characteristics
Chatbot	*D1*	**Primary Goal**	User support Action execution Coaching Data processor
	D2	**Knowledge Domain**	Specific Domain Restricted Domain
	D3	**Intelligence Framework**	Classic Rule-based AI Rule-based Retrieval Hybrid
	D4	**Integrated Service**	None Single Multiple
	D5	**Additional Human Support**	Present Not present
	D6	**Gamification**	Present Not present
	D7	**Service Provided**	Interpersonal Intrapersonal
	D8	**Socio-emotional Behaviour**	Present Not present
	D9	**Interface Personification**	Present Not present
	D10	**Front-end User Interface**	App Tool or Device

Perspective	Design Dimension		Characteristics
Chatbot-User Interaction	*D11*	**Communication Modality**	Only voice Multimodality
	D12	**Interaction Modality**	Graphical Interactive
	D13	**Lenght of Conversation**	Singol-turn Multi-turn
	D14	**Duration Single Interaction**	Short Interaction Medium-Long Interaction
	D15	**Leader of Conversation**	Chatbot-driven User-driven Mixed
	D16	**Frequency of Interactions**	Always When Required
	D17	**Number of Participants**	Individual Two or More
	D18	**Chatbot Role**	Facilitator Expert

D4 Service Integration refers to the Inter-agent classification criterion [7,23] and is intended to indicate whether the chatbot has the ability to offer additional third-party services (e.g., activate robots, place orders, manipulate GUIs, etc.). The identified features are divided into None, when a chatbot has no additional services beyond the one for which it is implemented (LARRI [42]), Single, when it is capable of performing only one additional service (Agroexpert [46]), and Multiple, when it provides two or more services (Xiadong [47]).

Dimension D5 Additional Human support analyzes whether the chatbot offers the possibility of contacting an external operator (human agent) for direct assistance or in circumstances where the chatbot is unable to provide an answer to the user's query. With the D6 Gamification design dimension proposed in [39], we want to analyze whether or not game elements (such as quizzes) are present in a generic chatbot to support users' learning or entertainment activities. Although these latter features can be considered on a par with a service offered by the chatbot and therefore included within the more generic D4 dimension, it was decided to distinguish these dimensions as potentially representing an interesting design element. In fact, although available case studies in manufacturing seem to suggest little use of such elements in the manufacturing domain, it is pointed out that in other domains such elements have some relevance, for example considering the healthcare domain for Additional Human Support [48,49] and the e-learning domain for Gamification [50,51].

D7 Service provided dimension indicates whether the chatbot falls within the user's personal domain and if it has user memory or not [7,20]. The first category of D7 are Static chatbots that deal with users by simply delivering the service and have no memory of the operators, such as Max [43] and Bot-X [45]. In contrast, Adaptive are those chatbots that have memory of the users and tasks they have previously performed, such as LARRI [42] and Chip [52].

Finally, through the design dimensions D8 Socio-emotional behavior, D9 Interface Personification lies the desire to analyze chatbots from the point of view of human similarity, i.e., the degree to which a user perceives his or her digital interlocutor to be similar to a human being [7]. Specifically, the D8 represents a synthesis of the dimensions Socio-emotional behavior and Personality processing/adaptability [38,39]. Its purpose is to indicate whether the chatbot has ability to show empathy. D9, on the other hand, is inspired by the Interface personification and avatar representation dimensions and aims to indicate whether the chatbot possesses virtual personification through a name and an avatar. Finally, D10 Front-end User Interface indicates whether it is developed as an App, and thus downloadable to various devices, or whether it is integrated directly into enterprise tools and devices.

Chatbot-User interaction perspective

First dimension identified to characterize the interaction between chatbot, and user is D11 Communication modality. This element refers to the architecture presented in section X and indicates whether the chatbot can receive input and/or respond through a single interaction channel (Text or Voice) or through multiple modalities (text, voice, video, etc.).

Dimension D12 Interaction Modality aims to classify a chatbot according to the type of interaction allowed by the software. Specifically, it subdivides chatbots with graphical interaction from chatbots with interactive interaction. In the former, interaction between user and chatbot occurs through text-buttons containing predefined choices. In the latter case, interaction can occur through free text, without restrictions on input.

D13 Length of Conversation dimension evaluates the total number of turns the chatbot considers to provide the response [4,27,40] Specifically, in Single-turn chatbots, the response is One-shot (e.g., Xiadong [47]), and for instance provided by considering only the user's current message. In Multi-turn chatbots instead, multiple iterations are considered to provide the response (e.g.,[42])

D14 Duration Single interaction indicates the average duration of a single interaction with the chatbot. This dimension takes as reference the dimensions relation duration and duration of interaction proposed by the reference taxonomies.
In addition, for the development of a chatbot it is necessary to set who is the conversation leader [4]. Specifically, D15 Leader of Conversation distinguishes conversational agents into User-driven, in which the user is the leader of the conversation and Chatbot-driven in which the chatbot leads the conversation and finally mixed solutions in which the leaders alternate.

Dimension D16 Frequency of Interactions distinguishes manufacturing chatbots into two categories and highlights the frequency in the use of the chatbot by users. The first indicates those chatbots used every time the operator needs to perform the task. The second refers to those chatbots used only when necessary.

D17 Number of Participants classifies the chatbot in relation to the number of possible participants during a single interaction with the conversational agent. Although cases analyzed reported 1:1 interaction, this dimension was still included in the taxonomy to emphasize the possibility of multiple interactions with the chatbot, for instance in a station with multiple operations and workers.

Finally, dimension D18 Chatbot Role indicates what kind of role the chatbot takes during the conversation. A chatbot may be classified as a facilitator if it facilitates the performance of the activity, or it may be classified as an expert if it transfers information that the operator has no knowledge of.

5. Extensive case studies analysis

To confirm and demonstrate how manufacturing chatbot case studied identified are distributed among characteristics an extensive analysis has been conducted. Each chatbot has been deeply investigated and mapped across the eighteen dimensions and forty-two characteristics. The authors have opted for two analyses: a first qualitative analysis of diffusion of each characteristic among the manufacturing case studies and a second analysis by parallel chart to show trends in the relationship between characteristics. For those cases where it was not possible to confidently identify a characteristic, a named characteristic "not available N/A" was added.

Qualitative Analysis

Table 2 shows the results achieved because of mapping each case with its characteristics. It is important to emphasize that since this is a small sample of observations, only a few qualitative hypotheses can be made, which should be properly validated through the classification of a larger sample.

However, the analysis carried out showed that there is a slight preference in developing conversational agents with the goal of providing assistance to operators when performing their tasks (40% User support). Furthermore, in line with the papers found on various scientific databases (Scopus, ResearchGate, Google Scholar), it is highlighted that the use of chatbots for the activation of robots or mechanical components, is still at an early stage of research. Another interesting result concerns the Intelligent Framework (D3) dimension. Findings showed that the Rule-based approach is the most widely used when it comes to conversational agents in manufacturing. Although an apparent balance of the characteristics of this dimension can be observed in Table 1, it is worth mentioning that AI Rule-based use Machine Learning techniques exclusively for the NLU module. Thus, it attests a slight trend to turn toward a classical, rule-based approach, although the use of ML techniques is not discouraged when useful for a better understanding of the operator's voice. Concerning knowledge domain (D2), the analysis also highlighted that the trend in manufacturing is to develop chatbots with unreduced degree of specialization. In fact, most studies identify themselves as chatbots with Restricted knowledge domain (60%). This means it is preferred to develop chatbots specialized in a certain area (or a set of activities or processes) rather than on a single, specific activity. Taking maintenance activities as an example, there has been a shift from chatbots such as LARRI [42] focused on assisting the activity of repairing a specific code (mechanical parts of an airplane), to more complex chatbots both capable of guiding operators in repair activities and assisting them in other processes. Examples include support in maintenance planning activities, process monitoring, and report writing [53].

Table 2 - Qualitative Analysis

Number of Chatbot used for classification = 20

	Design Dimension	Charateristics	Results	%
Chatbot Perspective	**D1 Primary Goal**	User support	8	40%
		Action execution	2	10%
		Coaching	5	25%
		Data processor	5	25%
	D2 Knowledge Domain	Restricted Domain	12	60%
		Specific Domain	8	40%
	D3 Intelligent Framework	Classic Rule-based	5	25%
		AI Rule-based	5	25%
		Retrieval	5	25%
		N/A	5	25%
	D4 Service Integration	None	9	45%
		Single	4	20%
		Multiple	7	35%
	D5 Additional Human Support	Present	1	5%
		Not present	19	95%
	D6 Gamification	Present	0	0%
		Not Present	20	100%
	D7 Service Provided	Interpersonal	11	55%
		Intrapersonal	6	30%
		N/A	3	15%
	D8 Socio-emotional Behaviour	Present	5	25%
		Not present	15	75%
	D9 Interface Personification	Present	5	25%
		Not present	15	75%
	D10 Front-end User Interface	App	8	40%
		Tool or Device	7	35%
		N/A	5	25%
Chatbot-User Interaction Perspective	**D11 Communication Modality**	Only voice	12	60%
		Multimodality	8	40%
	D12 Interaction Modality	Graphical		0%
		Interactive	20	100%
		Multi-turn	6	30%
		N/A	8	40%
	D14 Duration Single Interaction	Short Interaction	13	65%
		Medium-Long Interaction	7	35%
	D15 Leader of Conversation	Chatbot-driven	3	15%
		User-driven	14	70%
		Mixed	3	15%
	D16 Frequency of Interactions	Always	6	30%
		When Required	14	70%
	D17 Number of Participants	N/A	20	100%
	D18 Chatbot Role	Facilitator	9	45%
		Expert	11	55%

55% of the conversational agents analysed identify themselves as Interpersonal chatbots, service providers without the ability to store operator information. In addition, about 55% of the conversational agents are designed to activate at least one third-party service. As previously mentioned, regarding Gamification and Additional Human Support, cases analysed did not feature information concerning the presence or absence of these characteristics. The analysis also shows that most chatbots are developed with a low degree of "humanization," resulting in a low degree of interest in Human Similarity. Specifically, 75 percent of the observed conversational agents exhibit neither the ability to show empathy nor possess virtual personification. Regarding the mode of interaction with the user, there is a tendency to develop chatbots with a single channel of voice communication (60%), although multimodality solutions are not disdained. The analysis also allows characterizing the interaction between operators and conversational agents developed in manufacturing. There is a tendency to develop chatbots based on an interaction of short duration (65% Short), guided totally or partially by the operator rather than by the conversational agent (70% User-driven), which occurs in most cases when operators express the need to use the chatbot (70% When Required). These results would suggest that this technology is being used in industrial facilities as a valuable support tool that can be relied upon to retrieve relevant information rapidly.

Parallel Coordinates Chart Analysis

From an in-depth reading of the cases and because of the qualitative analysis, it was noted that two classes of chatbots can be distinguished in manufacturing. The first includes those agents designed to be a source of information for users and to build data storage. These are not necessarily tied to an operational activity. The second class includes conversational agents designed as tools to support operational activities. These may also include data storage.

The first class, called Operative Support, includes conversational agents with Primary Goal "User Support" and "Action Execution." In contrast, the second class, called Knowledge Source includes "Data Processor" and "Coaching."

The analysis conducted in this section aims to evaluate feature deviations between the two chatbot classes and assess whether there are distinctive feature patterns within each class. For a qualitative assessment, Parallel Coordinates Plots were used. Each chatbot is represented by a single curve passing through each dimension and indicating for each the chatbot's design feature. This graphical representation provides an opportunity to easily identify any recurring patterns, as the curves of the conversational agents will tend to overlap and create areas of higher density at common features [54]. To diversify the two classes, the color blue was assigned to represent the curves of Operative Support and the color red for those belonging to the Knowledge Source class. The graph is shown in Figure 4. This second-level analysis confirmed that there is no clear distinction in the design characteristics of conversational agents based on the purpose for which they are implemented. This result provides an indication of how, when deciding to implement a chatbot in manufacturing, there are no defined rules or standards. The choice is left to a functional analysis of the development team, turning out to be strictly dependent on the needs of the scenario to be implemented. A confirmation of this result is the comparison of the classes in terms of the Intelligent Framework dimension (D3). It is possible to observe a balance of approaches used for implementation for each class. Results observed here are to be considered interesting, as one would have expected more characterization of the conversational agent classes and more differentiation in terms of individual design features. When analyzing the differences between the two classes, it can be seen that Knowledge Source chatbots tend to be developed with a more extensive knowledge domain (80% Restricted), while there would seem to be a balance for the characteristics of the D2 dimension with regard to Operational Support. These results can be considered a rationale for the nature of Knowledge Sources. Indeed, it is natural to think that conversational agents, whose ultimate goal is to represent a source of information for operators, should be developed with a broader knowledge domain to provide support in various business contexts. Knowledge Source chatbots also show a tendency not to be programmed to provide third-party services (60% None), while for Operative Support ones there is a pattern of having at least one service (70%). On the other hand, it is interesting to note that Operative Support chatbots tend to play the role of facilitators (60%), while Knowledge Source seems to lean more toward the role of experts (70%). These results suggest an important hypothesis in relation to the nature and purpose for which the chatbot is implemented. In fact, it is safe to assume that chatbots designed to support operational activities primarily play the role of facilitators by offering in most cases functionality to activate third-party services useful in the execution and completion of respective tasks. Conversely, it is equally safe to assume that Knowledge Source chatbots, generally play the role of experts on a task, for whom access to third-party services is most often not necessary since they are designed to be large sources of information themselves. Another distinction between the two classes relates to the Leader of Conversation and Front-end User Interface dimensions. In fact, Knowledge Source conversational agents have a strong tendency to be developed through a User-driven approach (90%), while in the Operational Support cases there is an increase in the number of applications where the conversation is totally or partially guided by the chatbot.

As far as the technical solution for implementing the chatbot, Operative Support seems to show a tendency to be developed as stand-alone tools or devices (60%), while Knowledge Source tends to be developed more as applications that can be downloaded directly to various devices (60%). Again, these results could be justified by the nature of the two chatbot classes. In fact, it is reasonable to assume that Operational Support conversational agents guide the operator step-by-step in the execution and completion of their tasks and that, such software, are developed with an independent device placed near the workstation. On the other hand, as far as Knowledge Source class is involved, chatbots are often developed through an application that can be downloaded to one's devices to maximize accessibility, and that the user guides the conversation to directly obtain the information he or she needs. Finally, it is interesting to note that, in Knowledge Source systems, there is a tendency to show a recurring pattern of features. In Figure 4, it is possible to observe areas of curve overlap at dimensions D8 -D9 and especially between dimensions D10-D18. In contrast, Operational Support, although relationships between dimensions can be identified here as well, suggests an apparent absence of any recurring pattern.

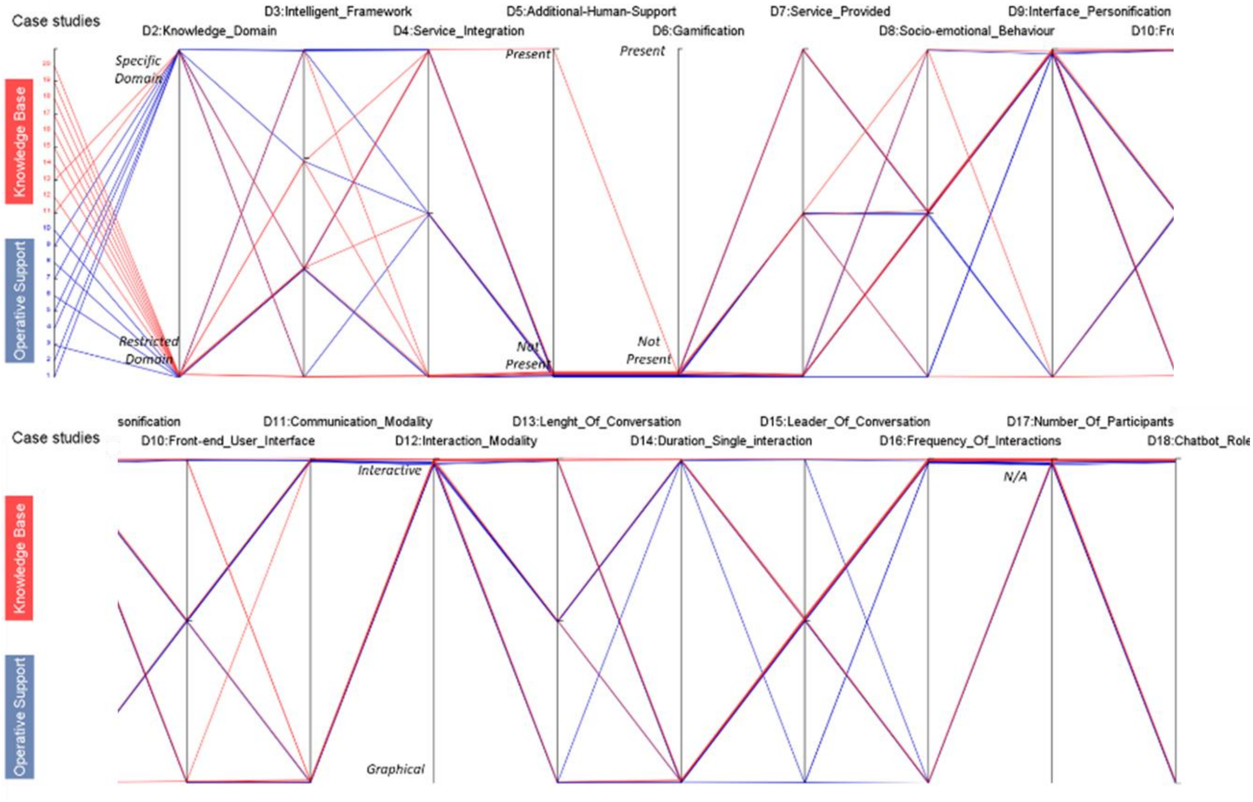


Figure 4 - Coordinates Parallel Chart

6. Conclusion

Conversational agents technology represents a simple, intuitive, and innovative solution that aims to revolutionize the field of Human-Machine Interaction in the manufacturing context. Although to date this technology has shown great potential and the various conversational agents have been developed in a variety of application areas, this technology is still in the early adopter's stage. This is especially evidenced by the absence of a reference standard and a general lack of mastery about their logical operation and characteristics. This is also reflected in the literature, in which conflicting statements about the classification criteria, general architecture and internal logic of operation of such systems are often found. This research analyzed the state of the art of the technology and proposed both a technical and functional guideline useful to organizations planning to adopt a conversational agent. From a technical point of view, a conceptual general architecture was developed to identify key development modules. Next, the focus was placed on the manufacturing sector. Here, conversational agents are configured as smart solutions applicable to various processes. These offer the potential to improve process performance by influencing user-satisfaction, reducing human error, and increasing the spread of business know-how. Such agents intervene both in alienating and repetitive operations and in hazardous operations where the operator needs to have hands and eyes free. Moreover, conversational agents could be an outsourced source of information for cybersecurity, with people asking support to avoid unwary behaviors [55]. The literature and analysis of application cases in manufacturing has shown the lack of common classification criteria and design features. In such a scenario, a reference taxonomy for conversational agents developed in manufacturing was developed following Nickerson's model. The taxonomy revealed important relationships among manufacturing chatbot design dimensions, bringing interesting insights to domain experts interested in manufacturing chatbot design.

Finally, as with all research, this work has some limitations, which offer opportunities for future research directions. While the authors thoroughly followed an established taxonomy development procedure [17], the limitations of this study mainly arise from the subjective choices inherent in any qualitative research approach. Notwithstanding, we applied a systematic empirical evaluation process and maintained a consistent unit of analysis throughout each case study investigated. In addition, such innovative topics often have few case studies to use to validate the research. This has made the application of our taxonomy to case studies limited. However, the authors consider the insights obtained an important step for more extensive analysis with new future manufacturing application cases. It is indeed expected that this technology will see an increase in application cases in the manufacturing context. In this regard, this work would serve as a tool for all partitioners to guide organizations toward greater understanding and adoption of such technology representative of beneficial Human-machine Interaction.

Appendix

Classes	Case study	Conversational Agent	Reference
Operative Support	1	Larri	[258]
	2	Max	[259]
	3	Bot-X	[260]
	4	Multi-modal	[261]
	5	Robot by voice	[262]
	6	Probot	[263]
	7	Ramp-Up	[264]
	8	CNC	[265]
	9	JAST	[266]
	10	Miaintenance	[267]
Knowledge Base	11	Xiaodong	[268]
	12	Training new employees	[269]
	13	FRASI	[270]
	14	MES	[271]
	15	Agriculture-Bot	[272]
	16	Agroexpert	[273]
	17	Transformer Mass-customization	[274]
	18	(Chip) Onboarding	[275]
	19	BPMN	[276]
	20	AECO industry	[277]

References

[1] A. Følstad, R. Halvorsrud, Communicating Service Offers in a Conversational User Interface: An Exploratory Study of User Preferences in Chatbot Interaction, ACM Int. Conf. Proceeding Ser. (2020) 671–676. https://doi.org/10.1145/3441000.3441046.

[2] E.H. Almansor, F.K. Hussain, Survey on Intelligent Chatbots: State-of-the-Art and Future Research Directions, Adv. Intell. Syst. Comput. 993 (2020) 534–543. https://doi.org/10.1007/978-3-030-22354-0_47.

[3] L.A. Flohr, S. Kalinke, A. Krüger, Di.P. Wallach, Chat or Tap? Comparing Chatbots with "Classic" Graphical User Interfaces for Mobile Interaction with Autonomous Mobility-on-Demand Systems, Proc. MobileHCI 2021 - ACM Int. Conf. Mob. Human-Computer Interact. Mob. Apart, MobileTogether. (2021). https://doi.org/10.1145/3447526.3472036.

[4] A. Følstad, P.B. Brandtzaeg, Chatbots and the New World of HCI, Interactions. 24 (2017) 38–42. https://doi.org/10.1145/3085558.

[5] M. McTear, Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots, Synth. Lect. Hum. Lang. Technol. 13 (2020) 1–251. https://doi.org/10.2200/S01060ED1V01Y202010HLT048.

[6] S. Syvänen, C. Valentini, Conversational agents in online organization–stakeholder interactions: a state-of-the-art analysis and implications for further research, J. Commun. Manag. 24 (2020) 339–362. https://doi.org/10.1108/JCOM-11-2019-0145/FULL/PDF.

[7] Q. Motger, X. Franch, J. Marco, Conversational Agents in Software Engineering: Survey, Taxonomy and Challenges; Conversational Agents in Software Engineering: Survey, Taxonomy and Challenges, (2021).

[8] D. Rooein, D. Bianchini, F. Leotta, M. Mecella, P. Paolini, B. Pernici, Chatting about processes in digital factories: A model-based approach, Lect. Notes Bus. Inf. Process. 387 LNBIP (2020) 70–84. https://doi.org/10.1007/978-3-030-49418-6_5/FIGURES/7.

[9] T.Y. Chen, Y.C. Chiu, N. Bi, R.T.H. Tsai, Multi-modal Chatbot in Intelligent Manufacturing, IEEE Access. 9 (2021). https://doi.org/10.1109/ACCESS.2021.3083518.

[10] D.H. Huang, H.E. Chueh, Chatbot usage intention analysis: Veterinary consultation, J. Innov. Knowl. 6 (2021) 135–144. https://doi.org/10.1016/J.JIK.2020.09.002.

[11] A.J.C. Trappey, C. V. Trappey, M.H. Chao, N.J. Hong, C.T. Wu, A vr-enabled chatbot supporting design and manufacturing of large and complex power transformers, Electron. 11 (2022). https://doi.org/10.3390/electronics11010087.

[12] S. Colabianchi, M. Bernabei, F. Costantino, Chatbot for training and assisting operators in inspecting containers in seaports, Transp. Res. Procedia. 64 (2022) 6–13. https://doi.org/10.1016/J.TRPRO.2022.09.002.

[13] D. Romero, J. Stahre, T. Wuest, O. Noran, P. Bernus, Å. Fast-Berglund, D. Gorecky, Towards an operator 4.0 typology: A human-centric perspective on the fourth industrial revolution technologies, CIE 2016 46th Int. Conf. Comput. Ind. Eng. (2016).

[14] F. Costantino, A. Falegnami, L. Fedele, M. Bernabei, S. Stabile, R. Bentivenga, New and emerging hazards for health and safety within digitalized manufacturing systems, Sustain. 13 (2021) 10948. https://doi.org/10.3390/su131910948.

[15] A. Adriaensen, F. Costantino, G. Di Gravio, R. Patriarca, Teaming with industrial cobots: A socio-technical perspective on safety analysis, Hum. Factors Ergon. Manuf. 32 (2022) 173–198. https://doi.org/10.1002/hfm.20939.

[16] R. Patriarca, A. Falegnami, F. Costantino, G. Di Gravio, A. De Nicola, M.L.M.L. Villani, WAx: An integrated conceptual framework for the analysis of cyber-socio-technical systems, Saf. Sci. 136 (2021) 105142. https://doi.org/10.1016/j.ssci.2020.105142.

[17] R.C. Nickerson, U. Varshney, J. Muntermann, A method for taxonomy development and its application in information systems, Eur. J. Inf. Syst. 22 (2013) 336–359. https://doi.org/10.1057/ejis.2012.26.

[18] A.M. Turing, Computing machinery and intelligence, Parsing Turing Test Philos. Methodol. Issues Quest Think. Comput. (2009) 23–65. https://doi.org/10.1007/978-1-4020-6710-5_3.

[19] J. Weizenbaum, ELIZA-A computer program for the study of natural language communication between man and machine, Commun. ACM. 9 (1966) 36–45. https://doi.org/10.1145/365153.365168.

[20] E. Adamopoulou, L. Moussiades, Chatbots: History, technology, and applications, Mach. Learn. with Appl. 2 (2020) 100006. https://doi.org/10.1016/J.MLWA.2020.100006.

[21] E. Quatrini, S. Colabianchi, F. Costantino, M. Tronci, Clustering Application for Condition-Based Maintenance in Time-Varying Processes: A Review Using Latent Dirichlet Allocation, Appl. Sci. 12 (2022). https://doi.org/10.3390/APP12020814.

[22] M. Bernabei, S. Colabianchi, F. Costantino, R. Patriarca, Using Natural Language Processing to uncover main topics in defect recognition literature, Proc. Summer Sch. Fr. Turco. (2021).

[23] E. Adamopoulou, L. Moussiades, An Overview of Chatbot Technology, Artif. Intell. Appl. Innov. 584 (2020) 373. https://doi.org/10.1007/978-3-030-49186-4_31.

[24] B. Luo, R.Y.K. Lau, C. Li, Y.W. Si, A critical review of state-of-the-art chatbot designs and applications, Wiley Interdiscip. Rev. Data Min. Knowl. Discov. 12 (2022) e1434. https://doi.org/10.1002/WIDM.1434.

[25] M. Serras, L. García-Sardiña, B. Simões, H. Álvarez, J. Arambarri, Dialogue Enhanced Extended Reality: Interactive System for the Operator 4.0, Appl. Sci. 2020, Vol. 10, Page 3960. 10 (2020) 3960. https://doi.org/10.3390/APP10113960.

[26] R. Agarwal, M. Wadhwa, Review of State-of-the-Art Design Techniques for Chatbots, SN Comput. Sci. 2020 15. 1 (2020) 1–12. https://doi.org/10.1007/S42979-020-00255-3.

[27] K. Ramesh, S. Ravishankaran, A. Joshi, K. Chandrasekaran, A survey of design techniques for conversational agents, Commun. Comput. Inf. Sci. 750 (2017) 336–350. https://doi.org/10.1007/978-981-10-6544-6_31.

[28] H. Chen, X. Liu, D. Yin, J. Tang, A Survey on Dialogue Systems, ACM SIGKDD Explor. Newsl. 19 (2017) 25–35. https://doi.org/10.1145/3166054.3166058.

[29] A. Khanna, B. Pandey, K. Vashishta, K. Kalia, B. Pradeepkumar, T. Das, A Study of Today's A.I. through Chatbots and Rediscovery of Machine Intelligence, Int. J. u-and e-Service. 8 (2015) 277–284. https://doi.org/10.14257/ijunesst.2015.8.7.28.

[30] Y. Wu, W. Wu, C. Xing, Z. Li, M. Zhou, Sequential Matching Network: A New Architecture for Multi-turn Response Selection in Retrieval-Based Chatbots, ACL 2017 - 55th Annu. Meet. Assoc. Comput. Linguist. Proc. Conf. (Long Pap. 1 (2017) 496–505. https://doi.org/10.18653/V1/P17-1046.

[31] S.A. Abdul-Kader, J. Woods, Survey on Chatbot Design Techniques in Speech Conversation Systems, IJACSA) Int. J. Adv. Comput. Sci. Appl. 6 (2015).

[32] B. Souvignier, A. Kellner, B. Rueber, H. Schramm, F. Seide, The thoughtful elephant: Strategies for spoken dialog systems, IEEE Trans. Speech Audio Process. 8 (2000) 51–62. https://doi.org/10.1109/89.817453.

[33] S. Quarteroni, Natural Language Processing for Industry: ELCA's experience, Informatik-Spektrum. 41 (2018) 105–112. https://doi.org/10.1007/s00287-018-1094-1.

[34] K. Greff, R.K. Srivastava, J. Koutnik, B.R. Steunebrink, J. Schmidhuber, LSTM: A Search Space Odyssey, IEEE Trans. Neural Networks Learn. Syst. 28 (2017) 2222–2232. https://doi.org/10.1109/TNNLS.2016.2582924.

[35] D. O'Shaughnessy, Invited paper: Automatic speech recognition: History, methods and challenges, Pattern Recognit. 41 (2008) 2965–2979. https://doi.org/10.1016/j.patcog.2008.05.008.

[36] J.G. Harms, P. Kucherbaev, A. Bozzon, G.J. Houben, Approaches for dialog management in conversational agents, IEEE Internet Comput. 23 (2019) 13–22. https://doi.org/10.1109/MIC.2018.2881519.

[37] Y. Ning, S. He, Z. Wu, C. Xing, L.J. Zhang, Review of deep learning based speech synthesis, Appl. Sci. 9 (2019). https://doi.org/10.3390/app9194050.

[38] A. Janssen, J. Passlick, D. Rodríguez Cardona, M.H. Breitner, Virtual Assistance in Any Context: A Taxonomy of Design Elements for Domain-Specific Chatbots, Bus. Inf. Syst. Eng. 62 (2020) 211–225. https://doi.org/10.1007/s12599-020-00644-1.

[39] M. Nißen, D. Selimi, A. Janssen, D.R. Cardona, M.H. Breitner, T. Kowatsch, F. von Wangenheim, See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons, Comput. Human Behav. 127 (2022). https://doi.org/10.1016/j.chb.2021.107043.

[40] S. Hussain, O. Ameri Sianaki, N. Ababneh, A Survey on Conversational Agents/Chatbots Classification and Design Techniques, Adv. Intell. Syst. Comput. 927 (2019) 946–956. https://doi.org/10.1007/978-3-030-15035-8_93.

[41] B. Schmidt, R. Borrison, A. Cohen, M. Dix, M. Gärtler, M. Hollender, B. Klöpper, S. Maczey, S. Siddharthan, Industrial virtual assistants - Challenges and opportunities, UbiComp/ISWC 2018 - Adjun. Proc. 2018 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput. Proc. 2018 ACM Int. Symp. Wearable Comput. (2018) 794–801. https://doi.org/10.1145/3267305.3274131.

[42] D. Bohus, A.I. Rudnicky, LARRI: A Language-Based Maintenance and Repair Assistant, (2005) 203–218. https://doi.org/10.1007/1-4020-3075-4_12.

[43] C. Li, J. Park, H. Kim, D. Chrysostomou, How can i help you? An intelligent virtual assistant for industrial robots, ACM/IEEE Int. Conf. Human-Robot Interact. (2021) 220–224. https://doi.org/10.1145/3434074.3447163.

[44] M. Casillo, F. Colace, L. Fabbri, M. Lombardi, A. Romano, D. Santaniello, Chatbot in industry 4.0: An approach for training new employees, Proc. 2020 IEEE Int. Conf. Teaching, Assessment, Learn. Eng. TALE 2020. (2020) 371–376. https://doi.org/10.1109/TALE48869.2020.9368339.

[45] C. Li, H.J. Yang, Bot-X: An AI-based virtual assistant for intelligent manufacturing, Multiagent Grid Syst. 17 (2021) 1–14.

[46] V. Nayak, P. R Nayak N, Sampoorna, Aishwarya, N.H. Sowmya, Agroxpert - Farmer assistant, Glob. Transitions Proc. 2 (2021) 506–512. https://doi.org/10.1016/J.GLTP.2021.08.016.

[47] J.S. Jwo, C.S. Lin, C.H. Lee, An Interactive Dashboard Using a Virtual Assistant for Visualizing Smart Manufacturing, Mob. Inf. Syst. 2021 (2021). https://doi.org/10.1155/2021/5578239.

[48] T. Kowatsch, M. Nißen, C.-H.I. Shih, D. Rüegger, D. Volland, A. Filler, F. Künzler, F. Barata, S. Haug, D. Büchter, B. Brogle, K. Heldt, P. Gindrat, N. Farpour-Lambert, & Dagmar L'allemand, Text-based Healthcare Chatbots Supporting Patient and Health Professional Teams: Preliminary Results of a Randomized Controlled Trial on Childhood Obesity, in: 7th Int. Conf. Intell. Virtual Agents (IVA 2017), 2017.

[49] P. Kucherbaev, A. Bozzon, G.J. Houben, Human-aided bots, IEEE Internet Comput. 22 (2018) 36–43. https://doi.org/10.1109/MIC.2018.252095348.

[50] A. Fadhil, A. Villafiorita, An adaptive learning with gamification & conversational UIs: The rise of CiboPoliBot, UMAP 2017 - Adjun. Publ. 25th Conf. User Model. Adapt. Pers. (2017) 408–412. https://doi.org/10.1145/3099023.3099112.

[51] I. Hidayatulloh, S. Pambudi, H.D. Surjono, T. Sukardiyono, N. Yogyakarta, Gamification on Chatbot-Based Learning Media: a Review and Challenges, Elinvo (Electronics, Informatics, Vocat. Educ. 6 (2021) 71–80. https://doi.org/10.21831/ELINVO.V6I1.43705.

[52] P. Chandar, Y. Khazaeni, M. Davis, M. Muller, M. Crasso, Q.V. Liao, N.S. Shami, W. Geyer, Leveraging Conversational Systems to Assists New Hires During Onboarding, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 10514 LNCS (2017) 381–391. https://doi.org/10.1007/978-3-319-67684-5_23.

[53] S. Wellsandta, Z. Rusak, S. Ruiz Arenas, D. Aschenbrenner, K.A. Hribernik, K.-D. Thoben, Concept of a Voice-Enabled Digital Assistant for Predictive Maintenance in Manufacturing, SSRN Electron. J. (2020). https://doi.org/10.2139/SSRN.3718008.

[54] J. Gonzalez, T. Dang, OutViz: Visualizing the Outliers of Multivariate Time Series, ACM Int. Conf. Proceeding Ser. (2021). https://doi.org/10.1145/3468784.3471606.

[55] A. Annarelli, S. Colabianchi, F. Nonino, G. Palombi, The Effectiveness of Outsourcing Cybersecurity Practices: A Study of the Italian Context, Lect. Notes Networks Syst. 360 LNNS (2022) 17–31. https://doi.org/10.1007/978-3-030-89912-7_2.

[56] J.N. Pires, Robot-by-voice: Experiments on commanding an industrial robot using the human voice, Ind. Rob. 32 (2005) 505–511. https://doi.org/10.1108/01439910510629244/FULL/XML.

[57] N. Ade, N. Quddus, T. Parker, S.C. Peres, ProBot – A Procedure Chatbot for Digital Procedural Adherence:, Https://Doi.Org/10.1177/1071181320641054. 64 (2021) 224–228. https://doi.org/10.1177/1071181320641054.

[58] M. Zimmer, A. Al-Yacoub, P. Ferreira, N. Lohse, Towards Human-Chatbot Interaction: A Virtual Assistant for the Ramp-up Process, UKRAS20 Conf. "Robots into Real World" Proc. 3 (2020) 108–110. https://doi.org/10.31256/QX5DT5V.

[59] F. Longo, A. Padovano, Voice-enabled Assistants of the Operator 4.0 in the Social Smart Factory: Prospective role and challenges for an advanced human–machine interaction, Manuf. Lett. 26 (2020) 12–16. https://doi.org/10.1016/J.MFGLET.2020.09.001.

[60] M.E. Foster, C. Matheson, Following Assembly Plans in Cooperative, Task-Based Human-Robot Dialogue, in: Conf. Proc. 12th Work. Semant. Pragmat. Dialogue (Londial 2008), 2008.

[61] A. Augello, G. Pilato, A. Machi, S. Gaglio, An approach to enhance chatbot semantic power and maintainability: Experiences within the FRASI project, Proc. - IEEE 6th Int. Conf. Semant. Comput. ICSC 2012. (2012) 186–193. https://doi.org/10.1109/ICSC.2012.26.

[62] S. Mantravadi, A.D. Jansson, C. Møller, User-Friendly MES Interfaces: Recommendations for an AI-Based Chatbot Assistance in Industry 4.0 Shop Floors, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 12034 LNAI (2020) 189–201. https://doi.org/10.1007/978-3-030-42058-1_16/FIGURES/6.

[63] M. Kiruthiga Devi, M.S. Divakar, V. Vimal Kumar, M.D.E. Jaincy, R.A. Kalpana, S.R.M. Kumar, FARMER'S assistant using ai voice bot, 2021 3rd Int. Conf. Signal Process. Commun. ICPSC 2021. (2021) 527–531. https://doi.org/10.1109/ICSPC51351.2021.9451760.

[64] F.C.T. Wu, O.N.J. Hong, A.J.C. Trappey, C. V. Trappey, VR-enabled chatbot system supporting transformer mass-customization services, Adv. Transdiscipl. Eng. 12 (2020) 291–300. https://doi.org/10.3233/ATDE200088.

[65] A. Lòpez, J. Sànchez-Ferreres, J. Carmona, L. Padrò, From Process Models to Chatbots, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 11483 LNCS (2019) 383–398. https://doi.org/10.1007/978-3-030-21290-2_24/FIGURES/6.

[66] M. Locatelli, E. Seghezzi, L. Pellegrini, L.C. Tagliabue, G.M. Di Giuda, Exploring natural language processing in construction and integration with building information modeling: A scientometric analysis, Buildings. 11 (2021). https://doi.org/10.3390/buildings11120583.

6.2. Appended Paper 4 Under Review: Enhancement of cybersecurity through digital intelligent assistant

Highlights

- A Digital Intelligent Assistant (DIA) for cybersecurity management is presented
- A structural design and functional features for DIA are presented
- The importance of positive and beneficial human-machine interaction is highlighted
- A RASA chatbot for a phishing attack scenario is developed

Abstract

The research proposes the application of digital intelligent assistants as proactive agents that can support employees in dealing with cybersecurity issues. Cyber attacks around the world are constantly increasing. Users are required to recall security procedures and rules. Moreover, attacks are constantly evolving and following different patterns. The study presents how a digital intelligent agent can backup agent during and after an attack. The application of digital intelligent assistance technology helps to reduce the cognitive load and pressure that users feel during downtime. In addition, the solution enhances attack reporting by decreasing the shame experienced by the victims. The research proposes a methodological design defining the agent's technical and functional characteristics. The solution is developed using the RASA framework and evaluated through a case study based on a phishing attack scenario. The introduction of this innovative technology in a workplace faced technical, social, and organizational challenges, showing benefits, limitations, and risks to all users.

Keywords: conversational system; natural language processing; human-AI collaboration; smart assistant; chatbot

1. Introduction

In recent years, many organizations have been the targets of cyber attacks and data breaches. Attackers all over the world are constantly developing new ways and strategies for breaking into and compromising even the most powerful security systems and gaining access to sensitive information (Annarelli et al., 2022; Linton et al., 2014). This has exposed not only company secrets but also the personal information of millions of consumers resulting in both economic and reputational loss for organizations. The latest CLUSIT report from October 2022 (CLUSIT, 2022) reported that attacks around the world have increased by 9 percent over the previous year and are getting more serious in quantity and sophistication. Moreover, according to the data available in the Common Vulnerabilities and Exposures (CVE) database, only in 2021 there have been reported around fifty-five vulnerabilities per day.

These data confirm that all organizations should stay aware and secure themselves to be able to detect a vulnerability and withstand an attack. However, the necessary equipment, personnel, time, and, most importantly, skills to face these challenges are rare (Palmer et al., 2016). Effective cybersecurity management requires organizations to invest in both new training programs that aim to increase cybersecurity awareness and the ability to detect an attack (Kweon et al., 2021). Nevertheless, IT and operational managers are also asked to be prepared to manage a possible disaster and reduce the risk and consequences of a cyber attack.

The increase of support people, training, and documentation are the most obvious options. However, the added costs associated with staff, software, building space, and so on make this option unfeasible for many organizations. In addition, when talking about cybersecurity, shame and reticence can have a strong impact. When workers realize that they have caused a cybersecurity incident they often feel guilt and shame, trying not to communicate the error and making the consequences of the attack worse (Renaud et al., 2021).

The scenario just described fits into a context in which organizations are experimenting with the Industry 5.0 paradigm. First, Industry 5.0 stresses the importance of adopting a human-centric approach to digital technologies including artificial intelligence applications (Nahavandi, 2019). There is a demand not to work on a machine vs. human project but to work on a partnership between them, seeing them as complementary and not competing. Second, Industry 5.0 stresses the importance of up-skilling and re-skilling the digital skills of workers by trying to close the digital skills gap for small and medium enterprises (Mukhuty et al., 2022). In this context, technology as a digital intelligent assistant can intervene in cybersecurity issues by fitting into a human-centric cybersecurity perspective, in line with the idea of industry 5.0.

It is believed that humans within the dynamics of a possible attack can be value added to the system. Humans with their human cognition, intuition, and flexibility can detect an anomaly by reporting it to the machine (Zimmermann & Renaud, 2019), which intervenes with technical expertise, assigning itself the most procedural and repetitive jobs. Specifically, the idea behind this research is to incorporate a digital intelligent assistant (DIA) into cybersecurity management to be a backup agent during and after an attack. DIAs constitute a cost-effective and scalable solution. DIA operates with users' natural language and it allows for a fast and on-demand response. As a result, they can answer questions from unskilled, and indeed skilled, information security employees (Dutta et al., 2018).

Such applications can help manage cybersecurity at several stages: from assisting users in detecting an attack to supporting them in the response and recovery phases. Furthermore, DIA can individually and contextually communicate on a one-to-many basis. This last aspect is critically important for maintaining a confidential relationship with the digital assistant by reducing the pressure workers feel when victimized by a cyber attack. Moreover, DIA provides promising pervasive and easy access to information and applications, offering an appropriate tool for human-centric approaches (Gartler & Schmidt, 2021).

In the literature, there is a growing interest in using digital assistants to support operators in critical and complex operations. Few studies, highly specific for particular scenarios, consider cybersecurity (Yoo & Cho, 2022). There are also some commercial applications such as cyber helpline chatbots (The Cyber Helpline, n.d.) or AI agents capable of detecting an attack (Prasad et al., 2021; Tagato et al., 2016) but which do not consider as part of their task's interaction with individuals. However, the employment of virtual agents in the context of SME cybersecurity is still scarce (Franco et al., 2020). Moreover, to the best of the authors' knowledge, there is insufficient literature in the context of the implementation of DIA in cybersecurity.

The research purpose is to develop an innovative technological solution to help organizations in the cybersecurity domain. Specifically, the research proposes a methodological design for the DIA defining its technical and functional characteristics. The developed DIA will take action when questioned by the user and support the user in managing the attack and post-attack by reducing his cognitive load and social pressure. The design is tested on a DIA to assist employees in the case of a phishing attack.

The remainder of the paper is organized as follows. The next section briefly overviews the current state of the art in the DIA applications and will underline the gaps related to the use of DIA in cybersecurity management. Section 3 presents our methodological design for the development of the DIA. Section 4 outlines the case study in cybersecurity management, while the fifth section presents the discussions and concluding remarks.

2. Related work and motivation

Digital assistants in the literature are also referred to as conversational agents, chatbots or voice-bots, voice assistants, intelligent agents, or virtual agents. DIAs are defined as those systems capable of holding a conversation with the user. In computing, there is a tendency to consider intelligent an agent that can exhibit rational behavior, showing reasoning skills. Natural language understanding (NLU), pattern recognition, and machine learning are some examples of activities that an agent can perform by exhibiting this capability (McTear, 2020). Nevertheless, there is no commonly accepted definition of what an intelligent agent is. An agent can detect its environment through sensors and act accordingly through actuators (Russel & Norvig, 1995). According to this definition, an agent can be seen as a system that is able to perceive the external environment by processing input (as strings of bits) and interacts by providing output that is consistent with the input received, using algorithms that guide the agent in choosing the action to be taken. Such systems have seen considerable evolution in recent years due to advances made in the field of Artificial Intelligence (AI), particularly in Natural Language Processing (NLP) and Machine Learning (ML) (Motger et al., 2021). One of the main application areas involves chatbots that support the user in daily tasks such as booking hotels or restaurants (Tiwari et al., 2023), up to the more innovative personal assistants (Siri, Cortana) and home-assistants (Alexa) (Athreya et al., 2018). Other popular application fields are e-commerce and financial activities (Cui et al., 2017). Interesting are the prototypes developed in the healthcare sector. In this field, agents have been developed to support patients and caregivers with applications such as prescribing medications, managing patients' personalized therapies, and software for controlling and monitoring vital signs (Fitzpatrick et al., 2017; Galetsi et al., 2022). The field of Education has also been subject to the influence of such technology, for instance with chatbots used to assist students during the learning process (Hien et al., 2018). More recently and in line with the idea of a more human-centric industry are the applications of digital assistants in the manufacturing sector. In this context, several applications are spreading with the aim of not only improving customer and user experience but also making their internal processes more efficient. These systems represent one of the enabling technologies for the transformation of an organization into a smart factory, emerging as new interfaces for human-machine communication. Examples can be found of agents supporting activities in maintenance (Wellsandt et al., 2022; Wellsandta et al., 2020), assembly (Chen et al., 2021), control of industrial devices and robots (Kalaiarassan et al., 2021; Li & Yang, 2021), or machine voice control (Longo & Padovano, 2020). Moreover, chatbot applications focused on user training (Casillo et al., 2020) and the onboarding of new operators (Chandar et al., 2017) are also relevant. The overall goal of these agents is to increase the degree of mobility, autonomy, and independence of the operators, creating a user-centered system in which the operators play the role of decision-makers, relying on such technologies to perform the most repetitive and alienating tasks (Rooein et al., 2020). Other advantages found are those related to the speed with which agents explore their knowledge domain and extract required answers, demanding little initial training (Li et al., 2021). With comparable aims, DIAs fit the few application cases featured for cybersecurity management. The contributions presented are quite specific case studies that present little methodological structure on the choices made for their development. One of the earliest works is (Gulenko, 2014), in which the authors developed a chatbot to train users on issues such as passwords, privacy, and secure browsing. The chatbot was developed using one of the first techniques proposed in the literature: artificial intelligence markup language. The chatbot traverses through a search tree to find the most useful output. The agent works effectively to help users in their learning path, including a humanization component that makes the conversation more engaging for the user. Nevertheless, the database is still limited and follows a strict pattern-matching approach to detect an answer. (Palmer et al., 2016) develop a cognitive cyber security system able to understand, learn and make decisions related to security issues. Its goal is to support the most experienced security analysts in the detection phase of an attack. The system is not designed with the logic of the digital assistant since it does not involve continuous interaction with humans, but it merely studies all variables and draws evidence-based conclusions. The research work of (Dutta et al., 2018) confirms the possible opportunities for applying the DIA technology in the cybersecurity domain, but it is still on a conceptual base without providing details for its development. More comprehensive is the work of (Franco et al., 2020) which combines neural networks and Natural Language Processing (NLP) with multiple cybersecurity aspects. The paper introduces a SecBot that can identify attacks during the conversation and can provide insights about risks and economic impacts. Their approach is technically advanced and it is oriented to a more skilled audience within the organization. A different technical approach adopted a broad knowledge base resulting in an apparent ability to respond quickly to multiple issues but with ultimately worse performance (Hamad & Yeferny, 2020). Recently, a cybersecurity chatbot was built specifically to support employees in the healthcare sector (Pears et al., 2021). In this work, RASA software is used, and performance is evaluated with a small group of respondents. The work is still in an embryonic stage but the performance to date was promising. The work proposed by (Yoo & Cho, 2022) is also promising. A telegram chatbot using Dialogflow detects an SNS phishing attack and provides suggestions to the victim. The chatbot is proactive and developed training in a convolutional neural network. Finally, two more recent research papers focused on issues such as increasing cyber awareness and training users through quizzes and chatbots (el Hajal et al., 2021; Fung & Lee, 2022). These works employed Google Dialogflow and Whatsapp. Their solutions are promising; however, they use a knowledge base that is useful to train the users before the attack and not support them during the attack.

The work proposed in this research fits into this growing trend of contributions by proposing DIA technology in the cybersecurity domain. The paper presents high-level, medium-level, and low-level goals. The high-level goal of the paper is to study how DIA can support cybersecurity management in an organizational context. The medium-level goal is to provide a structural design and functional features that make DIAs scalable and adaptable to multiple situations. The low-level goal is to contribute knowledge with the implementation of a real DIA capable to adapt the conversation depending on whether it is talking to an experienced or less skilled user. This solution is intended to be cost-effective and easy to use also for non-professionals. The DIA supports the user during and after an attack, it makes it understand the steps to follow and protect the individual and the business while providing useful guidance to increase user awareness. From a technical point of view, the agent is developed using the RASA framework and uses ML techniques for the NLU module.

At the core of its operation are the techniques of intents classification and slot filling (Adamopoulou & Moussiades, 2020; Gou et al., 2023). Using this approach, the agent can keep track of the information provided and query the user about missing information. Once all slots are filled, it can easily generate a response. This approach was chosen because a cyber attack is characterized by several parameters, and the joint consideration of those is needed to provide a suggestion on how to respond.

3. Proposed methodology

To realize the application described above, the authors followed a software engineering methodology (Sommerville, 2011) to identify relevant methodological design steps, decision-making options, information, benefits, and risks. The methodology consists of four macro steps.

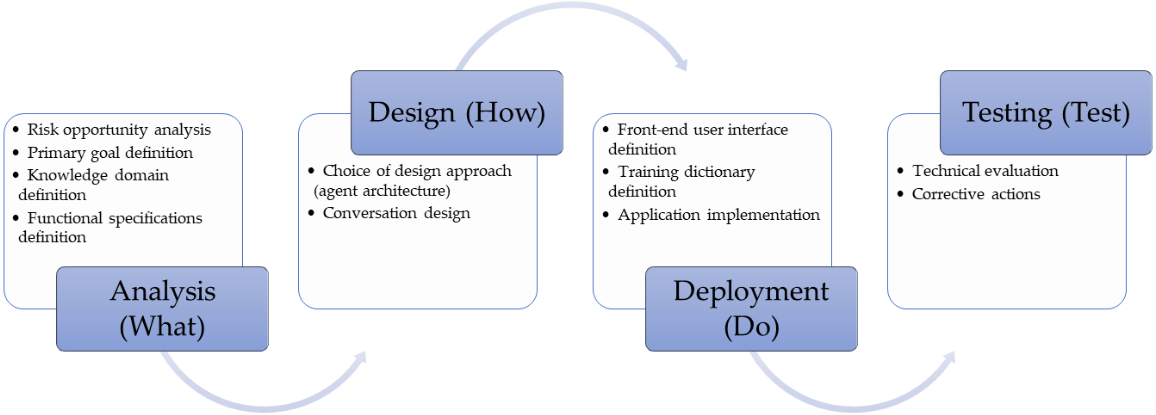

Risk opportunity analysis
Primary goal definition
Knowledge domain definition
Functional specifications definition
Analysis (What)
Design (How)
Choice of design approach (agent architecture)
Conversation design
Front-end user interface definition
Training dictionary definition
Application implementation
Deployment (Do)
Testing (Test)
Technical evaluation
Corrective actions

Figure 1 - Methodology steps

- Analysis (WHAT): The first key step is the Analysis (What) phase in which developers' efforts are focused on defining the functional specifications and application boundaries of the conversational agent.

The first step involves risk and opportunity analysis. This is used to justify the investment, the inclusion of the agent within the process, and define the boundaries of the operational activity in which it operates. The impact on workers must also be assessed at this stage. On the one hand, the inclusion of such technology improves user experience and user satisfaction; on the other hand, it can lead to risks such as cybersickness or workers' susceptibility to change which should be monitored (Følstad & Halvorsrud, 2020; Li et al., 2021). Process impact should also be considered. This refers to aspects such as direct support to workers, the reduction of human error, but also compatibility with other systems or workplace design.

Once the risk-opportunity analysis has been finalized, the next step is the requirements analysis in which it will be essential to identify the functional characteristics that the chatbot should exhibit once implemented. The first choice is to define in detail its knowledge domain. Specifically, all contexts and intents that the chatbot will have to manage are defined. There are no predefined rules; the software designer must choose whether want an agent with a specific knowledge domain (e.g., execution of a single process task) or an extended domain (e.g., execution of the entire process). Finally, the requirements and characteristics of the agent are defined. The taxonomy defined by (Nißen et al., 2022) will be used to define the dimensions and qualities that a chatbot should have. Following a systematic process that included 103 real-world chatbots, their study created a taxonomy of design features for domain-specific virtual agents. The design taxonomy distinguishes three levels of analysis viewpoint: i) twelve dimensions pertaining to chatbot profile, appearance, and intelligence; ii) seven dimensions pertaining to chatbot-user interaction; and iii) three dimensions pertaining to user viewpoint.

- Design (HOW): In the second step, the actual way the chatbot is programmed is defined. Choices regarding the software's operating rationale are made in this step. The first choice concerns the approach to be used for the technical development of the agent. Reference is made to the choices of natural language understanding and response generation. Specifically, technical requirements are chosen following the architecture proposed by (Colabianchi et al., 2022). In their work, the architecture consists of five modules: Speech to Text (STT), Natural Language Understanding (NLU), Dialog Manager (DM), Natural Language Generation (NLG), and Text-To-Speech (TTS). For each, design choices must be made such as, for instance, selecting between a rule-based or retrieval approach.

Finally, the last step is the design of the conversation. At this point, the flow of the conversation and the mapping of all intent and actions will be defined. Here, depending on the approach selected, one must define the rules and heuristics to manage the rule-based decision logic or all the data corpus and machine learning algorithms in the case of a Retrieval approach.

- Deployment (DO): In the third step, the technical choices for the actual development of the software are made. The first step is the choice of a front-end user interface. A choice must be made between an application on mobile, a tool to install, or a stand-alone device. In addition, a decision is made as to which type of interface one should have with the user, whether text or speech. Finally, the actual implementation of the defined architecture proceeds.

- Testing (TEST): In the last phase, the performance of the chatbot in terms of performance and meeting the functional requirements defined in the analysis and design phase will be evaluated. As far as the authors know there is no reference standard for digital agents. Recently, some authors have proposed as a reference model the Software Product Quality Model defined within the ISO/IEC 25010 standard (Motger et al., 2021). This standard defines the quality characteristics that must be considered when evaluating the properties of a software system. Specifically, seven Quality Characteristics were selected through ISO 25010: Functional Suitability, Performance Efficiency, Compatibility, Usability, Reliability, Security, and Maintainability.

4. Use Case

The case study conducted for this research is based on the process of assisting an employee victim of a phishing cyber attack. Phishing attacks fall within the social engineering family of attacks. The phrase "social engineering" refers to a range of techniques in which attackers exploit human channels to achieve their objectives. In this situation, hacking efforts increasingly focus on human weaknesses by processing their information rather than on software or hardware flaws (Mitnick et al., 2003). Specifically, in a phishing attack, victims are targeted via a link, which is often discovered in an e-mail or text message addressed to them. Once clicked, the link may include malware or initiate other messages asking for personal information from the victim. The attack pattern related to a famous phishing attack that hit Google Docs in May 2017 (Levin, 2017) involving Google Docs emails and document accesses was followed to define the case study.

- Analysis

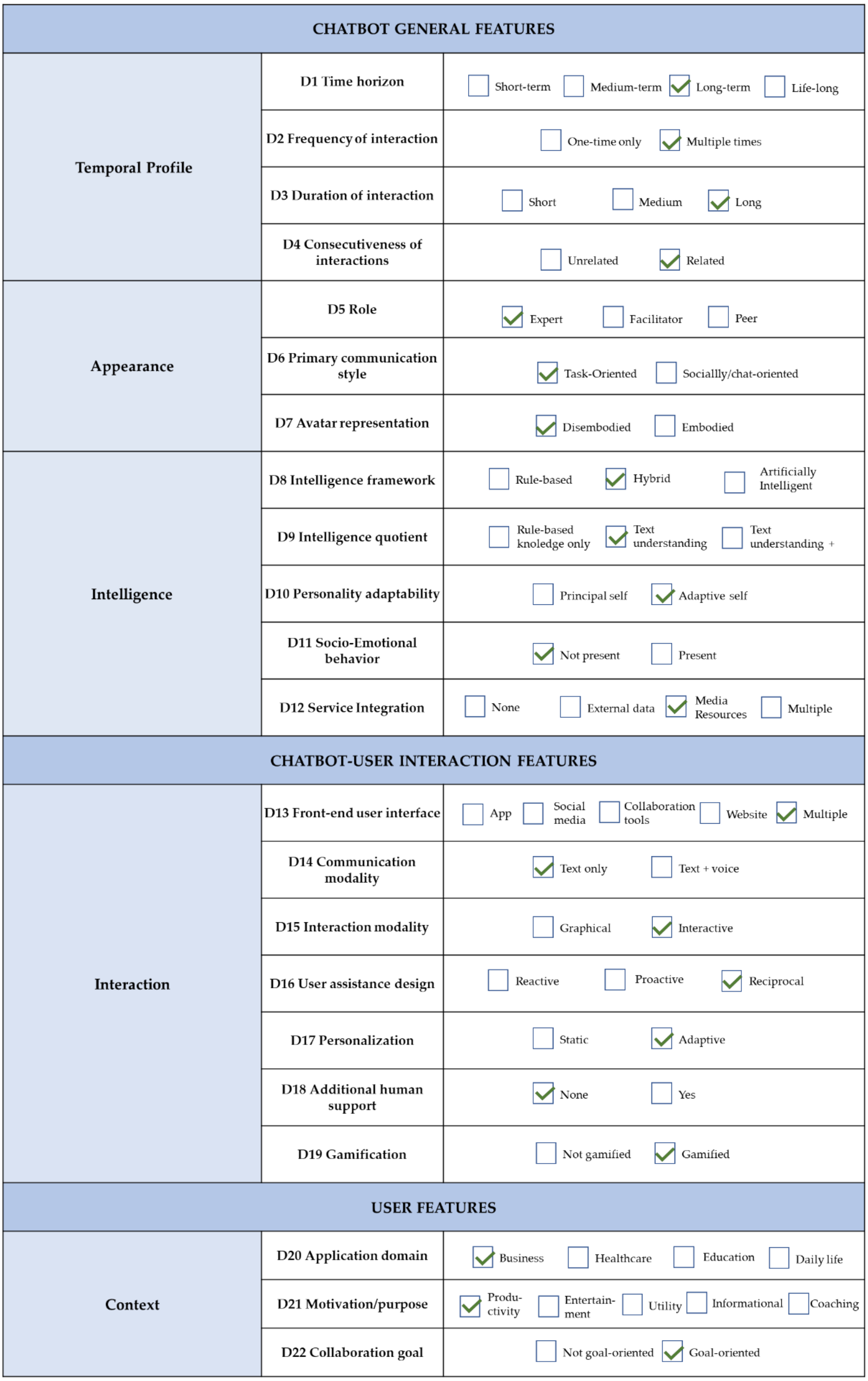
The digital intelligent assistant is defined as a business (D20), goal-oriented (D22), and expert (D5) agent for supporting employees in managing cybersecurity. The agent collaborates with the workers to accomplish a common task (D6) and it helps increase the productivity (D21) of organizations by improving the efficiency of resources (e.g., time, money, etc.).
The agent temporal profile is defined by a long-term (D1, D3) relationship characterized by multiple interactions (D2) over a certain period. The user indeed can interact multiple times with the agent to solve his/her security issue.
Due to the nature of the objective, the interactions are sequential and dependent on each other resulting in related consecutiveness of interaction (D4).
From a more technical point of view, the agent responses are generated on predefined rules and machine learning approaches resulting in a Hybrid Intelligent Framework (D8) and Text Understanding Intelligent Quotient (D9) enhanced by the integration of NLP techniques. The agent will not be integrated into other services in its first release (D12). The user can access the agent in multiple ways (app, web, etc.) from the workplace (D13). The communication will be text-only (D14) and let the user

express his/herself with free responses (D15). The leader of the conversation can be either the agent or the user (D16) and the agent can offer the user the possibility to contact a human agent in case of critical security situations (D18). Moreover, the agent can recognize whether a skilled or unskilled user is asking a question (D10) and can adapt and personalizes conversations based on user characteristics and conversation history (D17).

Finally, it was deemed unnecessary to establish a personification of the agent through an avatar representation (D7) or capabilities related to socio-emotion or empathic reactions to users' emotions (D11). In this case study, the agent will not integrate gamification elements (D19). Figure 2 summarizes all the functional specifications of the agent.

- Design

The design of the agent's architecture consists of three macro choices. First, for the NLU module, this agent will adopt NLP techniques for pre-processing the input text received by the user. Then, the user intent will be classified and mapped to the actions through a slot-filling approach (McTear, 2020). The text is then translated into values and data are transmitted to the dialog manager. It is the core module and manages the conversation with the user to achieve the goal of managing the issue and conducting the specified actions. The actions that the agent can take are of two types: internal actions, aimed at taking action to resolve a critical issue (e.g., terminate a connection to an application, request a password update, etc.); dialog actions aimed at continuing the conversation with the user to further understand the situation (Harms et al., 2019). Finally, a response is generated. For the case study proposed the design of the NLG module is based on an (AI) Rule-Based approach. Conversational agents in this category turn out to use machine learning techniques for the NLU module and are thus capable of handling more Intent than a simple Rule-Based, although they handle the NLG module through heuristics and rules like a Rule-Based (Adamopoulou & Moussiades, 2020). The answer is then translated into text and sent to the user. Figure 3 summarized the architecture of the agent. During this phase, the conversation between the agent and the user was also designed. Specifically, a flowchart representing the entire dialogue and all response paths was constructed so that rules for the rule-based approach could then be structured.



CHATBOT GENERAL FEATURES		
Temporal Profile	**D1 Time horizon**	□ Short-term □ Medium-term ☑ Long-term □ Life-long
	D2 Frequency of interaction	□ One-time only ☑ Multiple times
	D3 Duration of interaction	□ Short □ Medium ☑ Long
	D4 Consecutiveness of interactions	□ Unrelated ☑ Related
Appearance	**D5 Role**	☑ Expert □ Facilitator □ Peer
	D6 Primary communication style	☑ Task-Oriented □ Sociallly/chat-oriented
	D7 Avatar representation	☑ Disembodied □ Embodied
Intelligence	**D8 Intelligence framework**	□ Rule-based ☑ Hybrid □ Artificially Intelligent
	D9 Intelligence quotient	□ Rule-based knoledge only ☑ Text understanding □ Text understanding +
	D10 Personality adaptability	□ Principal self ☑ Adaptive self
	D11 Socio-Emotional behavior	☑ Not present □ Present
	D12 Service Integration	□ None □ External data ☑ Media Resources □ Multiple
CHATBOT-USER INTERACTION FEATURES		
Interaction	**D13 Front-end user interface**	□ App □ Social media □ Collaboration tools □ Website ☑ Multiple
	D14 Communication modality	☑ Text only □ Text + voice
	D15 Interaction modality	□ Graphical ☑ Interactive
	D16 User assistance design	□ Reactive □ Proactive ☑ Reciprocal
	D17 Personalization	□ Static ☑ Adaptive
	D18 Additional human support	☑ None □ Yes
	D19 Gamification	□ Not gamified ☑ Gamified
USER FEATURES		
Context	**D20 Application domain**	☑ Business □ Healthcare □ Education □ Daily life
	D21 Motivation/purpose	☑ Productivity □ Entertainment □ Utility □ Informational □ Coaching
	D22 Collaboration goal	□ Not goal-oriented ☑ Goal-oriented

Figure 2 - Cybersecurity Digital Intelligent Agent Functional specifications

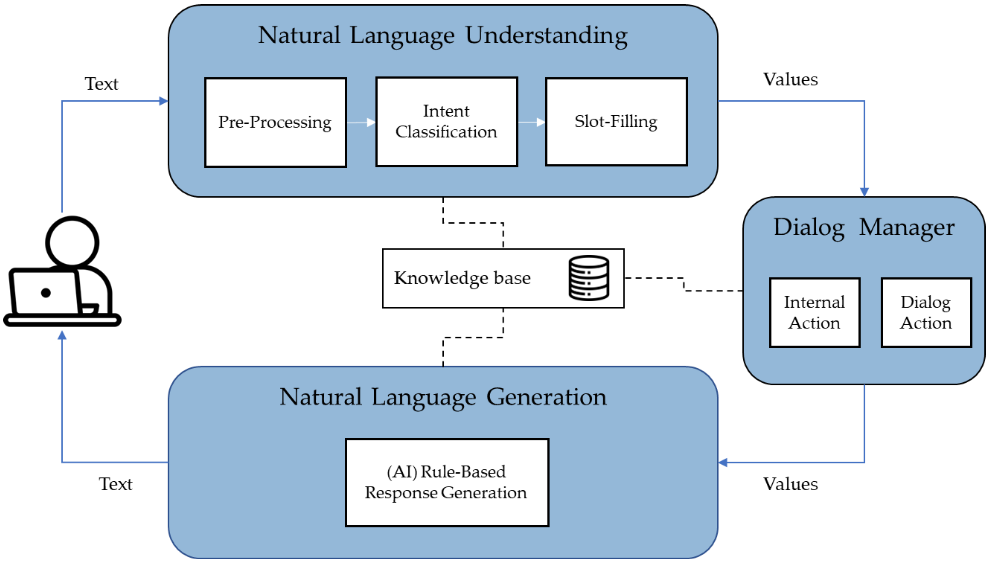


Figure 3 - Cybersecurity Digital Intelligent Agent Architecture

- Deployment

From a front-end user interface perspective, the agent will use a text-only conversation and will be accessed through a mobile or browser app from the workplace. In addition, the knowledge base of the agent and its training dictionary have been defined. The definition of the Training Dictionary is essential to train the software to identify the correct Intent and its slots but also to identify the vocabulary used by the users to decrease the Word Error Rate (Chiu et al., 2018).

The RASA framework was used for the actual deployment of the agent. Rasa is an open-source framework capable of handling text conversations through machine learning techniques. Rasa is structured on multiple classes that reflect the modules of the previously defined architecture. The classes defined are Natural Language Understanding, Domain, Rule, and Stories. In the Natural Language Understanding class, all intentions from a user are identified and stored. Intents and examples are used as training data for the model. In the Domain class, the actions that the chatbot can perform are defined. With the Rule class, parts of speech are described. Through rules, the Intent is mapped to a rule. In the Stories class, the agent is trained to respond correctly depending on what the user has mentioned in earlier stages of the conversation. Stories allow for a less inflexible conversation with the user. The slots are filled in no particular order; it is the agent who will be responsible for gathering all the information necessary to fill all the slots useful for generating a response. Figure 4 shows an example of the conversation. The example shows only an outline of the conversation. Specifically, the activation and filling out of a single form.

- Test

The evaluation of the innovative technology solution was done following the scheme proposed by (Motger et al., 2021). Their scheme is based on the ISO/IEC 25010 software product quality model. A qualitative evaluation of all features of the chatbot was conducted. A team of four experts was first defined for their evaluation. The team was structured as follows: a researcher and a university professor of industrial engineering, a software engineer, and a cybersecurity expert. The team, through an initial individual evaluation followed by a team discussion, evaluated the following quality characteristics: Functional suitability; Performance efficiency; Usability; Security. Concerning functional suitability, the team agreed to evaluate as effective both the content of the agent's knowledge base and the level of accuracy in interacting with the user. From a performance efficiency perspective, the solution is cost-effective and efficient in terms of resource utilization. Usability needs further evaluation. The discussion in the team states how the solution has potential in areas such as learnability, usability, and operability. However, it is difficult to assess at this stage the acceptance rate, interface accessibility, and user satisfaction for which a quantitative analysis with a user group is required. Finally, regarding security, features such as confidentiality and integrity are partially covered by basic security attention introduced. Certainly, when such a system is incorporated into a business environment, it is necessary to be prepared to manage related cyber risks. with techniques such as authentication (session) timeout and encryption (Shah & Panchal, 2022).

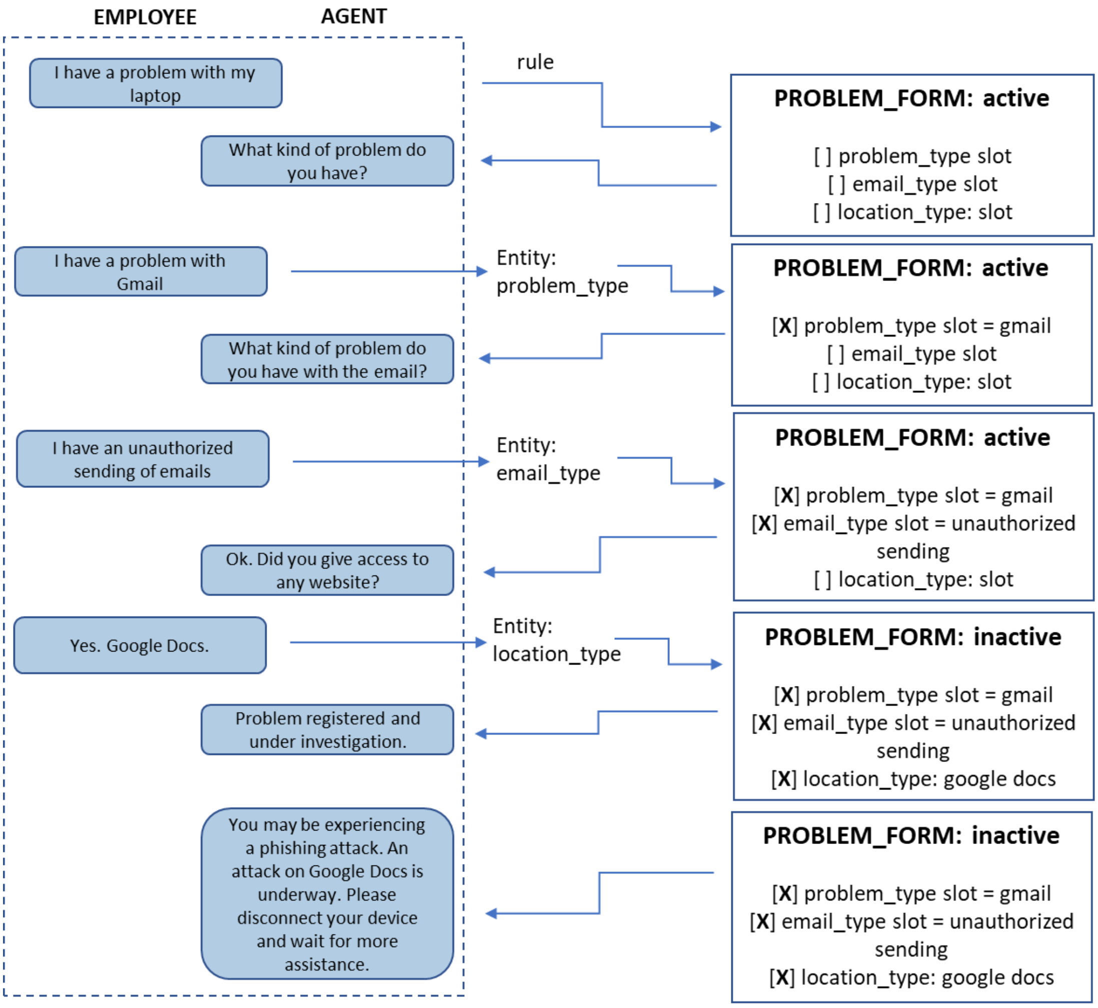

EMPLOYEE
AGENT
I have a problem with my laptop
rule
PROBLEM_FORM: active
[ ] problem_type slot
[ ] email_type slot
[ ] location_type: slot
What kind of problem do you have?
I have a problem with Gmail
Entity: problem_type
PROBLEM_FORM: active
[X] problem_type slot = gmail
[ ] email_type slot
[ ] location_type: slot
What kind of problem do you have with the email?
I have an unauthorized sending of emails
Entity: email_type
PROBLEM_FORM: active
[X] problem_type slot = gmail
[X] email_type slot = unauthorized sending
[ ] location_type: slot
Ok. Did you give access to any website?
Yes. Google Docs.
Entity: location_type
PROBLEM_FORM: inactive
[X] problem_type slot = gmail
[X] email_type slot = unauthorized sending
[X] location_type: google docs
Problem registered and under investigation.
You may be experiencing a phishing attack. An attack on Google Docs is underway. Please disconnect your device and wait for more assistance.
PROBLEM_FORM: inactive
[X] problem_type slot = gmail
[X] email_type slot = unauthorized sending
[X] location_type: google docs

Figure 4 - Example of conversation

5. Discussion

This article introduced an innovative DIA aimed at supporting employees during and after a cyber attack. The DIA developed has demonstrated success in improving the ability to respond to a cyber attack, particularly social engineering attacks that target human vulnerabilities. The tool helps reduce the pressure caused by any downtime related to the attack. In addition, the use of this tool reduces the shame of users who can seek help from a digital identity without having to immediately confirm that they have been the victims of the attack. Given the complex nature of cybersecurity issues, the solution also aims to reduce the cognitive load of workers who are often required to memorize multiple information, procedures, and rules.

However, the proposed agent also introduces organizational, technical, and social aspects that should not be underestimated. Most hybrid-augmented intelligence projects will rely on the goodwill of budget-controlling managers as a starting point. These managers must understand the significant benefits of incorporating a DIA into their operations, such as reduced costs and time, increased quality and performance, or any combination of these benefits (Cao et al., 2021). Then, managers would be able to defend their investments.

From a social point of view, we need to prepare employees for human-ai collaboration.

In the future, digital assistants could function as a type of digital collaborator. Employees must understand, and preferably experience, the strengths, problems, and risks of DIA to successfully integrate it into their systems. Otherwise, digital assistants may fail to meet expectations and lose or never regain the trust of employees. Data security and ethics, for example, are major risk areas. Consequently, non-technical solutions, such as training the workforce in human-AI collaboration, will be needed to overcome this obstacle. Furthermore, as the design of conversational AI for cybersecurity is in its infancy, it is important to design, develop, and evaluate effective designs to facilitate its implementation and integration with systems and people.

Finally, from an ethical perspective, the evolution of AI law and trustworthy AI introduces new legal requirements that AI-based solutions must meet. Our solution is in line with (European Commission, 2019), however anytime practitioners want to work on such a project, they should identify related gaps and study the operationalization of emerging AI laws and guidelines for the design, development, and ethical use of AI.

6. Conclusions and future work

In conclusion, the paper showed that the application of DIAs can support employees in dealing with cybersecurity issues. DIA supports employees by indicating the right procedure and assisting in performing actions according to security instructions. Real-time interaction with a virtual agent could strongly decrease the risk of cyber attacks. The paper confirms this opportunity, underlining the lack of applications in cybersecurity. The development of DIA in such a context requires the definition of an architecture and specific functional specifications. Currently, the new solution has been evaluated and validated in a laboratory environment. Future steps involve validating the solution and embedding it in a real work setting. Along with its introduction, the solution will be quantitatively evaluated by defining an evaluation protocol including interviews and questionnaires. Also, the way to conduct employee training incorporating this new innovative technology will be defined. Finally, future research includes expanding the agent's knowledge base by adding stories to handle more types of attacks and integrating the solution with other business systems to increase the agent's ability to act.

7. References

Adamopoulou, E., & Moussiades, L. (2020). An Overview of Chatbot Technology. Artificial Intelligence Applications and Innovations, 584, 373. https://doi.org/10.1007/978-3-030-49186-4_31

Annarelli, A., Colabianchi, S., Nonino, F., & Palombi, G. (2022). The Effectiveness of Outsourcing Cybersecurity Practices: A Study of the Italian Context. Lecture Notes in Networks and Systems, 360 LNNS, 17–31. https://doi.org/10.1007/978-3-030-89912-7_2

Athreya, R. G., Ngonga Ngomo, A. C., & Usbeck, R. (2018). Enhancing Community Interactions with Data-Driven Chatbots - The DBpedia Chatbot. The Web Conference 2018 - Companion of the World Wide Web Conference, WWW 2018, 143–146. https://doi.org/10.1145/3184558.3186964

Cao, G., Duan, Y., Edwards, J. S., & Dwivedi, Y. K. (2021). Understanding managers' attitudes and behavioral intentions towards using artificial intelligence for organizational decision-making. Technovation, 106, 102312. https://doi.org/10.1016/J.TECHNOVATION.2021.102312

Casillo, M., Colace, F., Fabbri, L., Lombardi, M., Romano, A., & Santaniello, D. (2020). Chatbot in industry 4.0: An approach for training new employees. Proceedings of 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, TALE 2020, 371–376. https://doi.org/10.1109/TALE48869.2020.9368339

Chandar, P., Khazaeni, Y., Davis, M., Muller, M., Crasso, M., Liao, Q. V., Shami, N. S., & Geyer, W. (2017). Leveraging Conversational Systems to Assists New Hires During Onboarding. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 10514 LNCS, 381–391. https://doi.org/10.1007/978-3-319-67684-5_23

Chen, T. Y., Chiu, Y. C., Bi, N., & Tsai, R. T. H. (2021). Multi-modal Chatbot in Intelligent Manufacturing. IEEE Access. https://doi.org/10.1109/ACCESS.2021.3083518

Chiu, C. C., Sainath, T. N., Wu, Y., Prabhavalkar, R., Nguyen, P., Chen, Z., Kannan, A., Weiss, R. J., Rao, K., Gonina, E., Jaitly, N., Li, B., Chorowski, J., & Bacchiani, M. (2018). State-of-the-Art Speech Recognition with Sequence-to-Sequence Models. ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing - Proceedings, 2018-April, 4774–4778. https://doi.org/10.1109/ICASSP.2018.8462105

CLUSIT. (2022). Rapporto Clusit Marzo 2022 – Clusit. https://clusit.it/rapporto-clusit/

Colabianchi, S., Bernabei, M., & Costantino, F. (2022). Chatbot for training and assisting operators in inspecting containers in seaports. Transportation Research Procedia, 64, 6–13. https://doi.org/10.1016/J.TRPRO.2022.09.002

Cui, L., Huang, S., Wei, F., Tan, C., Duan, C., & Zhou, M. (2017). Superagent: A customer service chatbot for E-commerce websites. ACL 2017 - 55th Annual Meeting of the Association for Computational Linguistics, Proceedings of System Demonstrations, 97–102. https://doi.org/10.18653/V1/P17-4017

Dutta, S., Joyce, G., & Brewer, J. (2018). Utilizing chatbots to increase the efficacy of information security practitioners. Advances in Intelligent Systems and Computing, 593, 237–243. https://doi.org/10.1007/978-3-319-60585-2_22

el Hajal, G., Abi Zeid Daou, R., & Ducq, Y. (2021). Human Firewall: Cyber Awareness using WhatApp AI Chatbot. 2021 IEEE 3rd International Multidisciplinary Conference on Engineering Technology, IMCET 2021, 66–70. https://doi.org/10.1109/IMCET53404.2021.9665642

European Commission. (2019). Ethics Guidelines for Trustworthy AI. https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): A randomized controlled trial. JMIR Mental Health, 4(2). https://doi.org/10.2196/MENTAL.7785

Følstad, A., & Halvorsrud, R. (2020). Communicating Service Offers in a Conversational User Interface: An Exploratory Study of User Preferences in Chatbot Interaction. ACM International Conference Proceeding Series, 671–676. https://doi.org/10.1145/3441000.3441046

Franco, M. F., Rodrigues, B., Scheid, J., Jacobs, A., Killer, C., Zambenedetti Granville, L., & Stiller, B. (2020). SecBot: a Business-Driven Conversational Agent for Cybersecurity Planning and Management. 16th International Conference on Network and Service Management, CNSM 2020, 2nd International Workshop on Analytics for Service and Application Management, AnServApp 2020 and 1st International Workshop on the Future Evolution of Internet Protocols, IPFutu.

Fung, Y. C., & Lee, L. K. (2022). A Chatbot for Promoting Cybersecurity Awareness. Lecture Notes in Networks and Systems, 370, 379–387. https://doi.org/10.1007/978-981-16-8664-1_33

Galetsi, P., Katsaliaki, K., & Kumar, S. (2022). Exploring benefits and ethical challenges in the rise of mHealth (mobile healthcare) technology for the common good: An analysis of mobile applications for health specialists. Technovation, 102598. https://doi.org/10.1016/J.TECHNOVATION.2022.102598

Gartler, M., & Schmidt, B. (2021). Practical challenges of virtual assistants and voice interfaces in industrial applications. Proceedings of the Annual Hawaii International Conference on System Sciences, 4063–4072.

Gou, Z., Li, Y., Liu, Y., & Gao, K. (2023). Topic model for personalized end-to-end task-oriented dialogue. Expert Systems with Applications, 212. https://doi.org/10.1016/J.ESWA.2022.118805

Gulenko, I. (2014). Chatbot for IT Security Training: Using Motivational Interviewing to Improve Security Behaviour. Undefined.

Hamad, S., & Yeferny, T. (2020). A Chatbot for Information Security. Undefined, 20(4).

Harms, J. G., Kucherbaev, P., Bozzon, A., & Houben, G. J. (2019). Approaches for dialog management in conversational agents. IEEE Internet Computing, 23(2), 13–22. https://doi.org/10.1109/MIC.2018.2881519

Hien, H. T., Cuong, P. N., Nam, L. N. H., Nhung, H. L. T. K., & Thang, L. D. (2018). Intelligent assistants in higher-education environments: The FIT-EBOt, a chatbot for administrative and learning support. ACM International Conference Proceeding Series, 69–76. https://doi.org/10.1145/3287921.3287937

Kalaiarassan, G., Franklin Prashanth, C., Prakash, M., & Phirke, S. (2021). Speech Recognition based Industrial Cloud Robot for Service-Oriented Sustainable Manufacturing. IOP Conference Series: Materials Science and Engineering, 1123(1), 012047. https://doi.org/10.1088/1757-899X/1123/1/012047

Kweon, E., Lee, H., Chai, S., & Yoo, K. (2021). The Utility of Information Security Training and Education on Cybersecurity Incidents: An empirical evidence. Information Systems Frontiers, 23(2), 361–373. https://doi.org/10.1007/s10796-019-09977-z

Levin, S. (2017). Google Docs users hit with sophisticated phishing attack in their inboxes | Hacking | The Guardian. https://www.theguardian.com/technology/2017/may/03/google-docs-phishing-attack-malware

Li, C., Park, J., Kim, H., & Chrysostomou, D. (2021). How can i help you? An intelligent virtual assistant for industrial robots. ACM/IEEE International Conference on Human-Robot Interaction, 220–224. https://doi.org/10.1145/3434074.3447163

Li, C., & Yang, H. J. (2021). Bot-X: An AI-based virtual assistant for intelligent manufacturing. Multiagent and Grid Systems, 17(1), 1–14. https://doi.org/10.3233/MGS-210340

Linton, J. D., Boyson, S., & Aje, J. (2014). The challenge of cyber supply chain security to research and practice – An introduction. Technovation, 34(7), 339–341. https://doi.org/10.1016/J.TECHNOVATION.2014.05.001

Longo, F., & Padovano, A. (2020). Voice-enabled Assistants of the Operator 4.0 in the Social Smart Factory: Prospective role and challenges for an advanced human–machine interaction. Manufacturing Letters, 26, 12–16. https://doi.org/10.1016/J.MFGLET.2020.09.001

McTear, M. (2020). Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots. Synthesis Lectures on Human Language Technologies, 13(3), 1–251. https://doi.org/10.2200/S01060ED1V01Y202010HLT048

Mitnick, K. D., Simon Foreword by Steve Wozniak, W. L., & Arynne, F. (2003). The art of deception: Controlling the human element of security. https://books.google.com/books?hl=en&lr=&id=rmvDDwAAQBAJ&oi=fnd&pg=PR7&ots=_fw-Px05U8&sig=-t9htwHF2coX5fdQjvfw82pd5Fw

Motger, Q., Franch, X., & Marco, J. (2021). Conversational Agents in Software Engineering: Survey, Taxonomy and Challenges; Conversational Agents in Software Engineering: Survey, Taxonomy and Challenges. https://arxiv.org/abs/2106.10901

Mukhuty, S., Upadhyay, A., & Rothwell, H. (2022). Strategic sustainable development of Industry 4.0 through the lens of social responsibility: The role of human resource practices. Business Strategy and the Environment, 31(5), 2068–2081. https://doi.org/10.1002/BSE.3008

Nahavandi, S. (2019). Industry 5.0-a human-centric solution. Sustainability (Switzerland), 11(16). https://doi.org/10.3390/SU11164371

Nißen, M., Selimi, D., Janssen, A., Cardona, D. R., Breitner, M. H., Kowatsch, T., & von Wangenheim, F. (2022). See you soon again, chatbot? A design taxonomy to characterize user-chatbot relationships with different time horizons. Computers in Human Behavior, 127. https://doi.org/10.1016/j.chb.2021.107043

Palmer, C., Angelelli, L., Linton, J., Singh, H., & Muresan, M. (2016). Cognitive cyber security assistants - Computationally deriving cyber intelligence and course of actions. AAAI Fall Symposium - Technical Report, FS-16-01-FS-16-05, 158–164.

Pears, M., Henderson, J., & Konstantinidis, S. T. (2021). Repurposing case-based learning to a conversational agent for healthcare cybersecurity. Public Health and Informatics: Proceedings of MIE 2021, 1066–1070. https://doi.org/10.3233/SHTI210348

Prasad, R. R., Rejimol Robinson, R. R., Thomas, C., & Balakrishnan, N. (2021). Evaluation of Strategic Decision taken by Autonomous Agent using Explainable AI. Proceedings of the 4th ISEA International Conference on Security and Privacy, ISEA-ISAP 2021. https://doi.org/10.1109/ISEA-ISAP54304.2021.9689715

Renaud, K., Searle, R., & Dupuis, M. (2021). Shame in Cyber Security: Effective Behavior Modification Tool or Counterproductive Foil? ACM International Conference Proceeding Series, 70–87. https://doi.org/10.1145/3498891.3498896

Rooein, D., Bianchini, D., Leotta, F., Mecella, M., Paolini, P., & Pernici, B. (2020). Chatting about processes in digital factories: A model-based approach. Lecture Notes in Business Information Processing, 387 LNBIP, 70–84. https://doi.org/10.1007/978-3-030-49418-6_5/FIGURES/7

Russel, S., & Norvig, P. (1995). Artificial Intelligence: A Modern Approach. Third Edition. PRENTICE HALL IN ARTIFICIAL INTELLIGENCE. https://zoo.cs.yale.edu/classes/cs470/materials/aima2010.pdf

Shah, M. H., & Panchal, M. (2022). Theoretical Evaluation of Securing Modules for Educational Chatbot. Proceedings - 2022 6th International Conference on Intelligent Computing and Control Systems, ICICCS 2022, 818–824. https://doi.org/10.1109/ICICCS53718.2022.9788120

Sommerville, I. (2011). Software engineering (9th ed). Pearson.

Tagato, H., Sakae, Y., Kida, K., & Asakura, T. (2016). Automated Security Intelligence (ASI) with auto detection of unknown cyber-attacks. NEC Technical Journal, 11(1), 45–48.

The Cyber Helpline. (n.d.). Retrieved October 24, 2022, from https://www.thecyberhelpline.com/

Tiwari, A., Khandwe, A., Saha, S., Ramnani, R., Maitra, A., & Sengupta, S. (2023). Towards personalized persuasive dialogue generation for adversarial task oriented dialogue setting. Expert Systems with Applications, 213. https://doi.org/10.1016/J.ESWA.2022.118775

Wellsandt, S., Klein, K., Hribernik, K., Lewandowski, M., Bousdekis, A., Mentzas, G., & Thoben, K. D. (2022). Hybrid-augmented intelligence in predictive maintenance with digital intelligent assistants. Annual Reviews in Control, 53, 382–390. https://doi.org/10.1016/J.ARCONTROL.2022.04.001

Wellsandta, S., Rusak, Z., Ruiz Arenas, S., Aschenbrenner, D., Hribernik, K. A., & Thoben, K.-D. (2020). Concept of a Voice-Enabled Digital Assistant for Predictive Maintenance in Manufacturing. SSRN Electronic Journal. https://doi.org/10.2139/SSRN.3718008

Yoo, J., & Cho, Y. (2022). ICSA: Intelligent chatbot security assistant using Text-CNN and multi-phase real-time defense against SNS phishing attacks. Expert Systems with Applications, 207. https://doi.org/10.1016/J.ESWA.2022.117893

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

7. Conclusions and future work

7.1. Conclusions

Digital transformation has changed the scenario in which organizations operate and will continue to do so in the coming years. Most critical resources, such as public services, healthcare, electricity, and transportation, are all online. And threat actors are aware of this. Taking a large supply chain or critical smart grid can inflict more damage than previous cyber attacks. As we have seen in this thesis, associated with such digital transformation are several challenges. Most of these involve the issue of cybersecurity and the growing cybersecurity workforce gap. There is a growing need to create a more cybersecurity-resilient society. Organizations cannot take the risk of falling victim to an attack that can bring substantial economic and image damage, as well as possible risks related to workers' safety.

While the need to grow the global cybersecurity workforce is raising, it is clear that no one organization, government, or institution can fill this gap alone. To make a real impact in preventing and reducing cyber attacks, as well as safeguarding the people they threaten, requires active and ongoing engagement and collaboration among companies, academics, and governments.

Much has been invested in cyber risk awareness campaigns in recent years [278], [279]. Awareness affects the entire organization, seeking to increase understanding of cyber threats and empower people to be both safer and aware of their importance in mitigating or avoiding incidents. Furthermore, in this process of raising awareness, it is emphasized the importance of integrating the three key dimensions and overlapping domains of people, processes, and technology in organizations [280]. A systemic approach to cybersecurity is desirable for organizations. From the people's perspective, a strong cybersecurity culture helps minimize the impact of human vulnerabilities and risks from employee behavior. In addition, organizations should constantly review and update their cybersecurity policies and processes to ensure that they still meet all requirements. Such changes, as seen in previous chapters, should be disseminated and communicated throughout the organization. Finally, these aspects must be integrated with technology. Although technology should on the one side be independent of people, on the other hand, people must interface with it and be aware of it. Moreover, the near future, following the Industry 5.0 paradigm, shows how increased performance can be achieved if investments are made in positive collaboration between humans and technology.

However, what has just been described is difficult to apply within organizations. In this process, there is still a struggle to understand how to make the human element part of the cyber resilience process [117]. First, as has been extensively presented in the thesis, it is important to manage the vulnerabilities generated by individuals with their behaviors and mistakes. Second, efforts are now being made through the most innovative training and technology not only to mitigate risks but also to leverage human factors as a contributor to cyber resilience. It is in this setting that this thesis work has been focused on. During the dissertation, it proceeded to investigate possible operational management improvements for cyber resilience. To do this, it was necessary to analyze the state of the art of research on the topic and identify promising areas of research that offered room for innovative developments.

Figure 14 maps the perspective and the results accomplished during the thesis through a block diagram.

After a general exploration of the topic of cyber resilience and the domains in which it is studied, the state-of-the-art analysis was narrowed down to the approaches available in literature to enhance cyber resilience. The extensive literature review answered *research question 1* showing how several contributions focused on systems security. However, the growing trend of contributions analyzing cyber incidents also from a safety perspective is observed, underscoring the paradigm shift taking place in the study of cybersecurity. In addition, the need to consider humans more in the cybersecurity loop emerged, broadening the vision to include the ability of organizations to withstand an attack and learn from it from a cyber resilience perspective. These concepts enabled the definition of the research direction followed in the second part of the thesis work. The research direction and linked research questions encompassed both managerial and technical issues in a cross-disciplinary dimension in line with the industrial and management doctoral path undertaken. The specific results of the second part of the thesis are:

- *About research question 2.* The first result was to characterize the individual interacting with processes and technology. Specifically, we wanted to identify and map the human factors involved in cybersecurity by identifying how they could be a driver or a barrier to increasing the cyber resilience of the organization.
- *About research question 3.* The second result was related to the first lever of intervention to support organizations. Specifically, it involved discussing the effectiveness of a set of organizational cybersecurity outsourcing practices and reasoning about the decision to proceed in-house or outsource cyber management.
- *About research question 4.* The third result was related to the second lever of intervention. In this case, it was intended to leverage human-machine interaction for cyber resilience. The technology chosen was conversational agents. The goal was both to develop a common architecture and taxonomy to describe them and to develop a prototype agent for cybersecurity.

Regarding the detailed three research outcomes formalized, some general conclusions may be drawn from the generated research contributions. In detail, and as summarized in the figure, this thesis project suggests the following.

The first contribution relates to the research presented in Chapter 4. Several articles were analyzed and human factors that characterize human behavior were outlined. For each, circumstances that make the factor a barrier or driver for cyber resilience were identified. These factors were then integrated with the NIST cybersecurity framework. Such integration helps and invites practitioners to also consider humans in cybersecurity risk assessment procedures and in prioritizing and achieving cyber resilience objectives. Moreover, the research pointed out that it is no longer possible to exclude humans and in particular the human factor, from the cyber paradigm. Although it is the most vulnerable element in the system, if cyber threats are to be addressed, human involvement is inevitable. Distrust, restriction, and control toward employees, based on the assumption that they can be malicious, foster a destructive organizational culture and cause employees not to work for organizational security and safety [117].

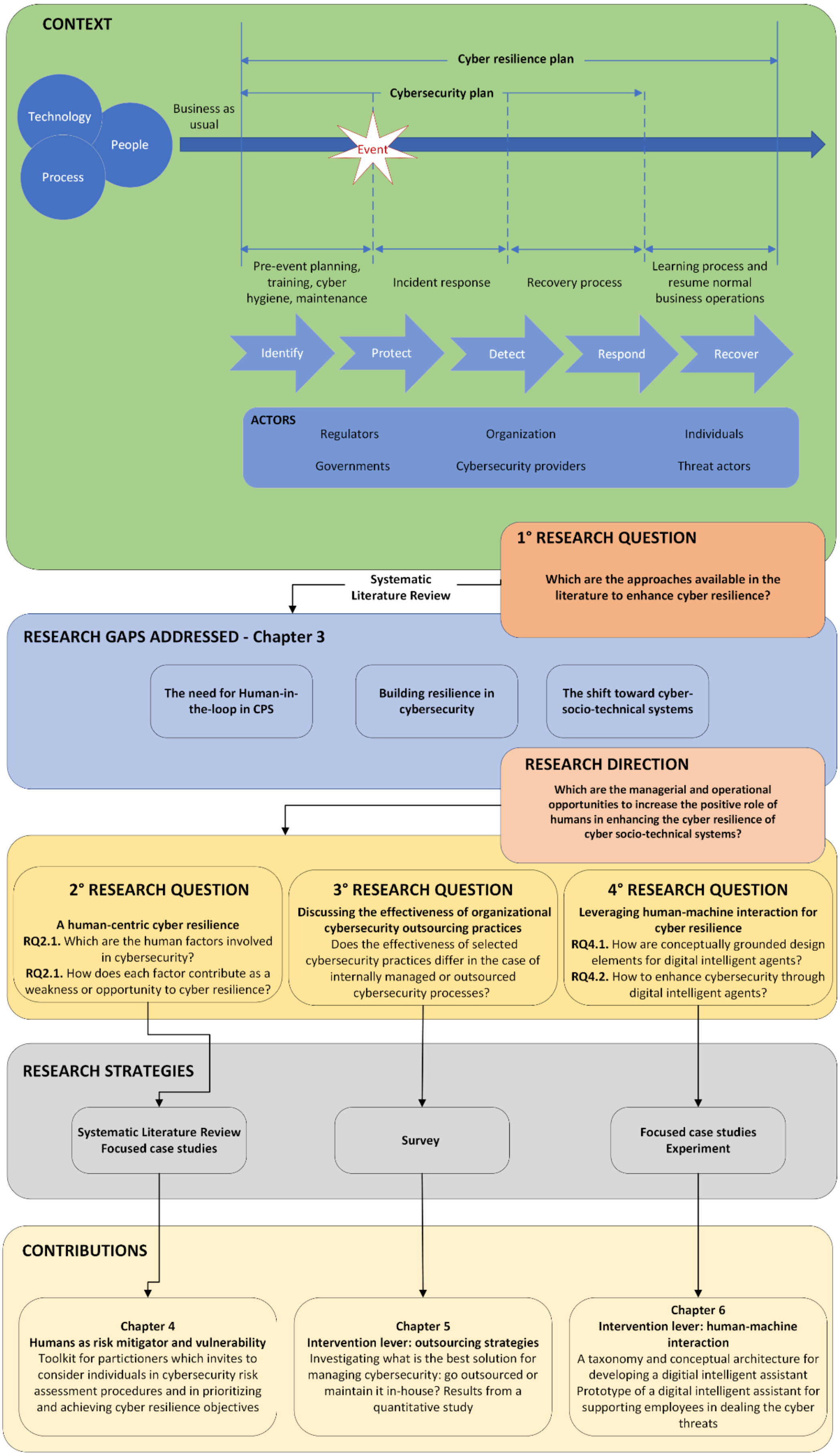
Instead, an organizational cyber resilience culture is needed in which people are willing to share responsibility and monitor the system. A cyber resilience culture is more than just cybersecurity awareness. It requires staff to know the security risks and the process for avoiding them. It is the construction and application of an operational business process that keeps the company cyber-safe and cyber-resilient.

An organization's business strategy should include an effective organization focused on information security. This means that organizations need policies that maintain a culture in which employees hesitate to bypass information security controls to accomplish their tasks [110]. Making humans protagonists, however, does not mean excluding technology, but working for a value-adding integration.

The second and third contributions focused on identifying two possible intervention levers that can support the individual in interacting with processes and technology. These levers work together with individuals allowing the generation of positive actions on the NIST cybersecurity framework functions.

Specifically, the work presented in Chapter 5 highlighted how cybersecurity is a complex and costly issue. Small and medium-sized enterprises struggle to have resources prepared for their management. Among the possible solutions is the use of external resources. This thesis surveyed a group of IT managers on their perceptions of the effectiveness of a selected group of organizational cybersecurity practices extracted from the categories of the NIST cybersecurity framework. The study achieves findings for the strand of research investigating what is the best solution for managing cybersecurity: go outsourced or maintain it in-house? Which one scales better in balancing risks and benefits? The data collected showed that small and medium-sized companies do not have many resources equipped to manage cybersecurity and that the process of training and awareness is in its beginning stages. Interesting considerations emerged regarding reputation and awareness of how major a cybercrime suffered is. Often, low awareness leads organizations to fail to report a cyber incident promptly. Organizations still fail to see the significant short- and long-term benefits this practice can bring. However, the trend, thanks partly to the new General Data Protection Regulation [58] and the NIS Directive [59] requiring organizations to notify the authority of data breaches and incidents, is growing. In addition, the analysis has shown that an awareness of the importance of lessons learned has increased. Organizations have realized that lesson learned sessions are critical to improving an organization's cyber resilience posture. They help assess incident response performance, identify challenges, and improve future response capabilities.

Finally, the third outcome concerns the development of a digital intelligent assistant to support organizations in managing cybersecurity. In the first phase of the research, the lack of a common vision of architecture and functional characteristics needed for the development of a digital assistant was identified. Therefore the first need was to develop a conceptual architecture useful to those who want to introduce a conversational agent into their systems. Next, a taxonomy of all the functionalities that an agent can have was developed. This taxonomy was then validated through case studies. Finally, a prototype digital intelligent assistant was developed to assist organizations in managing cybersecurity. The tool demonstrated how the inclusion of an enabling technology such as digital agents can contribute positively to the critical human factors identified in Chapter 4. The agent is a cost-effective solution capable of addressing issues such as shame, stress related to an attack, or the pressure and cognitive load associated with complex cybersecurity procedures.



1° RESEARCH QUESTION

Which are the approaches available in the literature to enhance cyber resilience?

RESEARCH GAPS ADDRESSED - Chapter 3

The need for Human-in-the-loop in CPS

Building resilience in cybersecurity

The shift toward cyber-socio-technical systems

RESEARCH DIRECTION

Which are the managerial and operational opportunities to increase the positive role of humans in enhancing the cyber resilience of cyber socio-technical systems?

2° RESEARCH QUESTION

A human-centric cyber resilience
RQ2.1. Which are the human factors involved in cybersecurity?
RQ2.1. How does each factor contribute as a weakness or opportunity to cyber resilience?

3° RESEARCH QUESTION

Discussing the effectiveness of organizational cybersecurity outsourcing practices
Does the effectiveness of selected cybersecurity practices differ in the case of internally managed or outsourced cybersecurity processes?

4° RESEARCH QUESTION

Leveraging human-machine interaction for cyber resilience
RQ4.1. How are conceptually grounded design elements for digital intelligent agents?
RQ4.2. How to enhance cybersecurity through digital intelligent agents?

RESEARCH STRATEGIES

Systematic Literature Review
Focused case studies

Survey

Focused case studies
Experiment

CONTRIBUTIONS

Chapter 4
Humans as risk mitigator and vulnerability
Toolkit for partictioners which invites to consider individuals in cybersecurity risk assessment procedures and in prioritizing and achieving cyber resilience objectives

Chapter 5
Intervention lever: outsourcing strategies
Investigating what is the best solution for managing cybersecurity: go outsourced or maintain it in-house? Results from a quantitative study

Chapter 6
Intervention lever: human-machine interaction
A taxonomy and conceptual architecture for developing a digitial intelligent assistant
Prototype of a digital intelligent assistant for supporting employees in dealing the cyber threats

Figure 14 - Thesis block diagram

7.2. Future research

The results presented in this thesis work also provided a direction for future research. The research provided a tool for practitioners useful in understanding how to integrate the human dimension within the NIST framework and benefit from it. This practice allows for a comprehensive risk analysis aimed at identifying best practices and areas of intervention to improve cyber resilience. Future research could include the creation of an artifact framework suitable for quantifying organizational resilience in cyber socio-technical systems. Such a solution will enable quantification of the effect of social attacks on cyber systems and offer tools to mitigate and/or guide human resources decisions. The goal is to reduce the consequences of risky human behaviors and to quantify the extent to which humans can be risk mitigators. In the framework, the agents involved are not limited to humans or organizations, but can also be artificial agents, such as the digital agents described in the thesis.

In addition, this research set the groundwork for a broader discourse moving towards a new paradigm of cybersecurity and cyber resilience. The rise of the concept of cyber resilience was discussed in the thesis. In addition, the human factor has long been debated as a positive influence on cyber management. The literature has shown that it is also due to individuals, so far considered the weak point of the system, and their vision and ability to move forward with successive adjustments that the system behaves resiliently. As noted these concepts characterize resilience engineering and its key ideas related to the new paradigm for safety management: "Safety-II”. Such a statement sets the basis for further reflection. Given the scenario that organizations face today and the complexity of cyber socio-technical systems, it is no longer possible to reason from an almost exclusive perspective of identifying, protecting, and detecting. It is not possible to continue treating the system as divided into parts, investing heavily in technical systems, lengthy, repetitive, and complex procedures, and blaming incidents on the human factor. This was all characteristic of early cybersecurity. Today we are working from a cyber resilience perspective. Cyber resilience seems to encompass cybersecurity, inviting practitioners to study the whole system from an integrated perspective. Being aware that every action is not negligible. The safe operation of the system is given by the continuous adaptation of humans and technology able to mitigate and recover post-incident. This all reminds of the shift from Safety-I to Safety-II in resilience engineering. However, such a shift toward a cyber resilience theory that includes humans and organizational elements and goes beyond controls and policies toward a perspective based on adaptive systems capabilities is still limited. These considerations at the end of the thesis pose additional questions that open up new research. First, following this logic, it may be stated that cyber resilience represents an integration of the concepts of cyber security and resilience engineering. The question then arises whether the research is moving towards a broader definition of cyber resilience that increasingly integrates the concepts of cybersecurity and resilience engineering. Next, we question whether we are in presence of a paradigm shift similar to that between Safety-I and Safety-II even for cybersecurity. Is there an on-going evolution from a Security-I to a Security-II, just as it was for safety? To date, the literature is still far from affirming such a paradigm shift. However, the growing number of attacks and their related consequences indicate that traditional cybersecurity is no longer sufficient. Moreover, the growing need for resilient performance, the increased complexity, and the advantages of a "human as a solution" perspective suggest that the emergence of Security-II in a more cyber resilience-oriented research agenda is possible.

8. Thesis publications

This paragraph provides the bibliographic details of the appended papers. The appended papers have been published in scientific peer-reviewed journals and conferences and are the outcomes of collaborations with one or more co-authors. The full-length papers are included in the corresponding paragraphs. Furthermore, several related manuscripts published in journals, conference proceedings, and other sources are also listed in 8.2. These manuscripts report other work conducted during these three years of PhD. Specifically, these researches were conducted using approaches, models, and methods explored in depth during the thesis work and reapplied to other research strands or application domains.

8.1. Appended Papers

Following a thesis conceptual order:

I. Colabianchi, S., Costantino, F., Di Gravio, G., Nonino, F., & Patriarca, R. (2021). Discussing resilience in the context of cyber physical systems. *Computers & Industrial Engineering*, *160*, 107534.

II. Annarelli, A., Colabianchi, S., Nonino, F., & Palombi, G. (2021, November). The Effectiveness of Outsourcing Cybersecurity Practices: A Study of the Italian Context. In *Proceedings of the Future Technologies Conference* (pp. 17-31). Springer, Cham.

III. *Under Review*: Colabianchi, S., Tedeschi A., Costantino F. (2022). Human-technology integration with industrial conversational agents: a conceptual architecture and a taxonomy for manufacturing. *Journal of Industrial Information Integration*

IV. *Under Review:* Colabianchi, S., Costantino, F. (2022). Enhancement of cybersecurity through digital intelligent assistant. *IEEE Access*

8.2. Related documents

- The following are two papers that used natural language processing and clustering techniques. The detail of these techniques is given in Chapter 3 of this thesis. The papers make references to two different application cases: a review of Condition-Based Maintenance techniques and an exploratory review of defect recognition topics in literature.

I. Quatrini, E., Colabianchi, S., Costantino, F., & Tronci, M. (2022). Clustering Application for Condition-Based Maintenance in Time-Varying Processes: A Review Using Latent Dirichlet Allocation. *Applied Sciences,* 12(2), 814.

Abstract: In the field of industrial process monitoring, scholars and practitioners are increasing interest in time-varying processes, where different phases are implemented within an unknown time frame. The measurement of process parameters could inform about the health state of the production assets, or products, but only if the measured parameters are coupled with the specific phase identification. A combination of values could be common for one phase and uncommon for another phase; thus, the same combination of values shows a high or low probability depending on the specific phase. The automatic identification of the production phase usually relies on clustering techniques. This is largely due to the difficulty of finding training fault data for supervised models. With these two considerations in mind, this contribution proposes the Latent Dirichlet Allocation as a natural language-processing technique for reviewing the topic of clustering applied in time-varying contexts, in the maintenance field. Thus, the paper presents this innovative methodology to analyze this specific research fields, presenting the step-by-step application and its results, with an overview of the theme.

II. Bernabei M., Colabianchi S., Costantino F., Patriarca R. (2021). Using Natural Language Processing to uncover main topics in defect recognition literature, *Paper presented at the Proceedings of the Summer School Francesco Turco*

Abstract: The issue of defect detection is particularly important namely in plant engineering, where it is crucial to ensure high-quality production by minimizing the number of defective parts. In the last years, the interest in the subject has grown a lot and the methods and approaches proposed for defect recognition are multiple. Therefore, when dealing with defect recognition researchers are faced with an increasing number of articles that slows them down in identifying the set of articles of their interest. This work aims to provide a baseline classification of articles based on emerging issues such as the investigated material, the production typology in which the material is included, and the type of analysis to be effected. For these reasons, the paper proposes an automatic solution based on text mining techniques. Specifically, the study applies Natural Language Processing (NLP) to articles' titles, abstracts, and keywords using two different approaches: K-Means clustering algorithm and Latent Dirichlet Allocation (LDA). K-Means is used to cluster the collection of documents into related groups based on the contents of the particular documents. LDA instead is used to classify the papers using the concept of topic modeling. Articles have been collected from Scopus database. The scope of the research is limited to journal and conference articles, published in English excluding articles classified as reviews, as well as book chapters, books, notes, erratum.

- Presented below is a systematic literature review that used Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) also used for the review of this thesis and presented in Chapter 3. Specifically in this paper, PRISMA was used to analyze the actions and strategies taken by supply chains to be resilient during pandemic disruption.

III. Bernabei, M., Colabianchi, S., & Costantino, F. (2022). Actions and Strategies for Coronavirus to Ensure Supply Chain Resilience: A Systemic Review. *Sustainability,* 14(20), 13243.

Abstract: The COVID-19 outbreak adversely impacted agri-food supply chains and caused a severe socio-economic crisis worldwide. Preventive measures taken by several countries have affected production and distribution. Moreover, producers have had to face difficulties related to changes in local and international export markets, a decrease in the labor force due to the spread of the virus, and challenges in harvesting, processing, and shipment of products. However, despite the extraordinary nature of the disruption, supply chains have demonstrated a fair, resilient, and sustainable crisis recovery. Although a large number of papers deal with supply chains and the pandemic's impact, a review of measures implemented that comprehensively includes resilience dimensions is still lacking. The scope of this paper is to survey available literature in order to understand whether there are classes of actions and strategies undertaken by meat supply chains in managing the pandemic. Documents were reviewed through a protocol based on the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) review technique. The survey highlights which actions have enabled supply chain resilience by underling virtuous behaviors and lessons learned. These findings support the need for further investigation of supply chain resilience and offer practitioners guidance toward a greater understanding of impacts and implementable strategies.

- In addition, cyber socio technical systems were introduced in the thesis work. These systems are a complex adaptive system in which social and human artifacts are inextricably intertwined with technical artifacts. Such systems are found in multiple domains. Two papers are presented below in which this concept has been applied to distribution centers.

IV. Bernabei, M., Colabianchi, S., Costantino, F., & Falegnami, A. (2022). Warehouse resilience framework for the Covid-19 disruption. *In Proceedings of 22nd International Working Seminar on Production Economics.*

Abstract: The COVID 19 pandemic consistently impacted Supply Chains (SCs), involving the agri-food sector. Compared to other SCs, agri-food SCs have not suffered complete interruptions, but the effect of various exogenous and endogenous phenomena has impacted all the actors in the chains. Within the agri-food SC, wholesalers play a strategic role influencing the performance and service level of a large number of players. Despite their importance, there has been no in-depth analysis of the effect of the pandemic on wholesalers. The paper fills this gap by proposing a framework to guide wholesaler during a severe disruption event. The MARLIN framework (fraMework wArehouse ResiLience dIstruptioN) starts from a collection of the principal warehouse KPIs and from the identification through a literature analysis of the most significant factors and indicators on the pandemic disruption. The proposed framework supports wholesalers in identifying the most critical warehouse areas and defining interventions to mitigate the effects of future phases of the disruption. The framework has been tested on a case study involving an Italian wholesaler's warehouse located in central Italy. The results obtained have demonstrated the effectiveness of the framework by highlighting aspects that are difficult to identify in an emergency situation.

V. *Under Review:* Bernabei, M., Colabianchi, S., Costantino, F., Romano, E., & Falegnami, A. (2022). A cyber-socio-technical system method for warehouse management in the face of perturbations. A case study on Covid-19. *International Journal of Management Science and Engineering Management*

Abstract: Endogenous and exogenous phenomena to the supply chains (SC) elements echo in variations of operations characteristic parameters. For wholesale warehouses, various disruptive events impact infrastructure or performance undermining the economic-organizational survival of the warehouse. For example, the Covid-19 impacted retail good SC, requiring modifications on their normal behaviour to adapt. Distributions of warehouses play a fundamental role, influencing performance and service level for all the SC players. Considering models and theories of Resilience

Engineering, this paper presents the MARLIN method (Method wArehouse ResiLience dIstruptioN) to identify areas to be focused on and mitigation actions to be implemented in the face of perturbations. The method has been tested on a case study involving an Italian warehouse proving to be effective in identifying intervention areas and disruption related KPIs. As a side benefit, MARLIN brought along a convenient general-purpose list of indicators relative to the different areas belonging to a typical warehouse.

- Finally, a second application of a digital intelligent assistant is presented below. In this paper, the theory detailed in Chapter 6 of this thesis is applied to develop an assistant that guides port operators in performing high safety risk operations.

VI. Colabianchi, S., Bernabei, M., Costantino, F. (2022). Chatbot for training and assisting operators in inspecting containers in seaports. *Transportation Research Procedia, 64, 6-13.*

Abstract: The paper presents the chatbot applicability for the health and safety of workers in the container transportation context. Starting from a literature review of risks and hazardous activities in sea container terminals, the paper underlines the need of innovative systems to ensure the lowest level of risks for labours. An analysis of the 4.0 technologies solutions in sea container terminals shows the lack of empirical application of chatbots in such a context. Focus is given to the current chatbot applications, and on the conceptual methodology for the chatbot design, defining five models and presenting a taxonomy for the chatbot feature definition. A case study shows the possible application of the conceptual methodology and the taxonomy, introducing the Popeye chatbot, consisting of a voice service, spoken language understanding component and an image processing app, to cope with the hazards in the process of examining freight and containers in dock areas. The main application of Popeye is the training of new employees involved in container safety-critical quality inspection and controls operations.

References

[1] F. Björck, M. Henkel, J. Stirna, and J. Zdravkovic, “Cyber resilience – Fundamentals for a definition,” in *Advances in Intelligent Systems and Computing*, 2015, vol. 353, pp. 311–316. doi: 10.1007/978-3-319-16486-1_31.

[2] A. Corallo, M. Lazoi, and M. Lezzi, “Cybersecurity in the context of industry 4.0: A structured classification of critical assets and business impacts,” *Comput Ind*, vol. 114, 2020, doi: 10.1016/j.compind.2019.103165.

[3] S. Colabianchi, F. Costantino, G. di Gravio, F. Nonino, and R. Patriarca, “Discussing resilience in the context of cyber physical systems,” *Comput Ind Eng*, p. 107534, Jul. 2021, doi: 10.1016/J.CIE.2021.107534.

[4] R. Langner, “Stuxnet: Dissecting a cyberwarfare weapon,” *IEEE Secur Priv*, vol. 9, no. 3, pp. 49–51, May 2011, doi: 10.1109/MSP.2011.67.

[5] D. Kushner, “The real story of stuxnet,” *IEEE Spectr*, vol. 50, no. 3, pp. 48–53, 2013, doi: 10.1109/MSPEC.2013.6471059.

[6] Z. Zheng and A. L. Narasimha Reddy, “Towards improving data validity of cyber-physical systems through path redundancy,” 2017. doi: 10.1145/3055186.3055189.

[7] Y. Hashimoto *et al.*, “Safety securing approach against cyber-attacks for process control system,” *Comput Chem Eng*, vol. 57, pp. 181–186, Oct. 2013, doi: 10.1016/J.COMPCHEMENG.2013.04.019.

[8] R. A. Ramadan, B. W. Aboshosha, J. S. Alshudukhi, A. J. Alzahrani, A. El-Sayed, and M. M. Dessouky, “Cybersecurity and Countermeasures at the Time of Pandemic,” *J Adv Transp*, vol. 2021, 2021, doi: 10.1155/2021/6627264.

[9] T. Weil and S. Murugesan, “IT Risk and Resilience-Cybersecurity Response to COVID-19,” *IT Prof*, vol. 22, no. 3, pp. 4–10, 2020, doi: 10.1109/MITP.2020.2988330.

[10] L. Ntasis, K. Koronios, and T. Pappas, “The impact of COVID-19 on the technology sector: The case of TATA Consultancy Services,” *Strategic Change*, vol. 30, no. 2, pp. 137–144, Mar. 2021, doi: 10.1002/JSC.2397.

[11] L. Kappelman *et al.*, “The 2020 SIM IT Issues and Trends Study,” *MIS Quarterly Executive*, vol. 20, no. 1, pp. 69–107, 2021.

[12] A. Garcia-Perez, M. P. Sallos, and P. Tiwasing, “Dimensions of cybersecurity performance and crisis response in critical infrastructure organisations: an intellectual capital perspective,” *Journal of Intellectual Capital*, 2021, doi: 10.1108/JIC-06-2021-0166.

[13] J. Pattison, “From defence to offence: The ethics of private cybersecurity,” *European Journal of International Security*, vol. 5, no. 2, pp. 233–254, Jun. 2020, doi: 10.1017/EIS.2020.6.

[14] ENISA, “Definition of Cybersecurity - Gaps and overlaps in standardisation — ENISA,” 2015. Accessed: Oct. 15, 2022. [Online]. Available: https://www.enisa.europa.eu/publications/definition-of-cybersecurity

[15] G. Locke and P. D. Gallagher, “Managing Information Security Risk: Organization, Mission, and Information System View.” 2011. Accessed: Oct. 15, 2022. [Online]. Available: https://www.nist.gov/publications/managing-information-security-risk-organization-mission-and-information-system-view

[16] ISO/IEC 27000, “Information technology-Security techniques-Information security management systems-Overview and vocabulary ISO/IEC 27000:2018(E).” Accessed: Oct. 15, 2022. [Online]. Available: www.iso.org

[17] A. Dorri, S. S. Kanhere, R. Jurdak, and P. Gauravaram, “Blockchain for IoT security and privacy: The case study of a smart home,” *2017 IEEE International Conference on Pervasive Computing and Communications Workshops, PerCom Workshops 2017*, pp. 618–623, May 2017, doi: 10.1109/PERCOMW.2017.7917634.

[18] G. Dhillon, R. Syed, and F. de Sá-Soares, “Information security concerns in IT outsourcing: Identifying (in) congruence between clients and vendors,” *Information and Management*, vol. 54, no. 4, pp. 452–464, Jun. 2017, doi: 10.1016/j.im.2016.10.002.

[19] C. M. Gutierrez and W. Jeffrey, “FIPS PUB 200 Minimum Security Requirements for Federal Information and Information Systems,” 2006.

[20] CLUSIT, “Rapporto Clusit Ottobre 2022– Clusit,” 2022. Accessed: Nov. 24, 2022. [Online]. Available: https://clusit.it/rapporto-clusit/#form_edl

[21] MITRE, “Browse cve vulnerabilities by date,” 2022. https://www.cvedetails.com/browse-by-date.php (accessed Nov. 24, 2022).

[22] Q. Sun, K. Zhang, and Y. Shi, “Resilient Model Predictive Control of Cyber-Physical Systems under DoS Attacks,” *IEEE Trans Industr Inform*, 2020, doi: 10.1109/TII.2019.2963294.

[23] J. Liu, X. Lu, and J. Wang, “Resilience Analysis of DC Microgrids Under Denial of Service Threats,” *IEEE Transactions on Power Systems*, 2019, doi: 10.1109/TPWRS.2019.2897499.

[24] R. Ma, P. Shi, Z. Wang, and L. Wu, “Resilient filtering for cyber-physical systems under denial-of-service attacks,” *International Journal of Robust and Nonlinear Control*, 2020, doi: 10.1002/rnc.4845.

[25] I. Tomić, M. Breza, and J. A. McCann, “Jamming-resilient control and communication framework for cyber physical systems,” 2019. doi: 10.1049/cp.2019.0132.

[26] Y. Guan and X. Ge, “Distributed Attack Detection and Secure Estimation of Networked Cyber-Physical Systems Against False Data Injection Attacks and Jamming Attacks,” *IEEE Trans Signal Inf Process Netw*, 2018, doi: 10.1109/TSIPN.2017.2749959.

[27] D. Senejohnny, P. Tesi, and C. de Persis, “A jamming-resilient algorithm for self-triggered network coordination,” *IEEE Trans Control Netw Syst*, 2018, doi: 10.1109/TCNS.2017.2668901.

[28] A. Ameli, A. Hooshyar, E. F. El-Saadany, and A. M. Youssef, “An Intrusion Detection Method for Line Current Differential Relays,” *IEEE Transactions on Information Forensics and Security*, 2020, doi: 10.1109/TIFS.2019.2916331.

[29] O. M. Anubi, C. Konstantinou, C. A. Wong, and S. Vedula, “Multi-model resilient observer under false data injection attacks,” 2020. doi: 10.1109/CCTA41146.2020.9206284.

[30] I. Jovanov and M. Pajic, "Relaxing Integrity Requirements for Attack-Resilient Cyber-Physical Systems," *IEEE Trans Automat Contr*, 2019, doi: 10.1109/TAC.2019.2898510.

[31] K. Paridari, N. O'Mahony, A. El-Din Mady, R. Chabukswar, M. Boubekeur, and H. Sandberg, "A framework for attack-resilient industrial control systems: Attack detection and controller reconfiguration," *Proceedings of the IEEE*, 2018, doi: 10.1109/JPROC.2017.2725482.

[32] S. Kim, Y. Won, I.-H. Park, Y. Eun, and K.-J. Park, "Cyber-Physical Vulnerability Analysis of Communication-Based Train Control," *IEEE Internet Things J*, 2019, doi: 10.1109/JIOT.2019.2919066.

[33] K. Yang *et al.*, "Enhanced resilient sensor attack detection using fusion interval and measurement history," 2018. doi: 10.1109/CODESISSS.2018.8525941.

[34] X. Zhou *et al.*, "Evaluating Resilience of Grid Load Predictions under Stealthy Adversarial Attacks," 2019. doi: 10.1109/RWS47064.2019.8971816.

[35] T. A. McDermott, "A rigorous system engineering process for resilient cyber-physical systems design," 2019. doi: 10.1109/ISSE46696.2019.8984569.

[36] M. Pajic, J. Weimer, N. Bezzo, O. Sokolsky, G. J. Pappas, and I. Lee, "Design and Implementation of Attack-Resilient Cyberphysical Systems: With a Focus on Attack-Resilient State Estimators," *IEEE Control Syst*, 2017, doi: 10.1109/MCS.2016.2643239.

[37] S. Wedaj, K. Paul, and V. J. Ribeiro, "DADS: Decentralized attestation for device swarms," *ACM Transactions on Privacy and Security*, 2019, doi: 10.1145/3325822.

[38] V. Bhavsar, A. Kadlak, and S. Sharma, "Study on Phishing Attacks," *Article in International Journal of Computer Applications*, vol. 182, no. 33, pp. 975–8887, 2018, doi: 10.5120/ijca2018918286.

[39] S. Mishra and D. Soni, "DSmishSMS-A System to Detect Smishing SMS," *Neural Comput Appl*, 2021, doi: 10.1007/S00521-021-06305-Y.

[40] F. L. Greitzer, J. R. Strozer, S. Cohen, A. P. Moore, D. Mundie, and J. Cowley, "Analysis of unintentional insider threats deriving from social engineering exploits," *Proc IEEE Symp Secur Priv*, vol. 2014-Janua, pp. 236–250, Nov. 2014, doi: 10.1109/SPW.2014.39.

[41] B. Walker, C. S. Holling, S. R. Carpenter, and A. Kinzig, "Resilience, adaptability and transformability in social-ecological systems," *Ecology and Society*, vol. 9, no. 2, 2004, doi: 10.5751/ES-00650-090205.

[42] S. Hosseini, K. Barker, and J. E. Ramirez-Marquez, "A review of definitions and measures of system resilience," *Reliab Eng Syst Saf*, vol. 145, pp. 47–61, Jan. 2016, doi: 10.1016/j.ress.2015.08.006.

[43] T. J. Vogus and K. M. Sutcliffe, "Organizational resilience: Towards a theory and research agenda," *Conf Proc IEEE Int Conf Syst Man Cybern*, pp. 3418–3422, 2007, doi: 10.1109/ICSMC.2007.4414160.

[44] R. Martin, "Regional economic resilience, hysteresis and recessionary shocks," *J Econ Geogr*, vol. 12, no. 1, pp. 1–32, Jan. 2012, doi: 10.1093/jeg/lbr019.

[45] E. Hollnagel, J. Pariès, D. Woods, and J. Wreathall, *Resilience engineering in practice: A guidebook*. 2011.

[46] E. Hollnagel, D. D. Woods, and N. Leveson, “Resilience engineering: Concepts and precepts,” in *Resilience Engineering: Concepts and Precepts*, vol. 15, no. 6, Ashgate Publishing Ltd, 2012, pp. 1–397. doi: 10.1136/qshc.2006.018390.

[47] R. Patriarca, A. Falegnami, F. Costantino, G. di Gravio, A. de Nicola, and M. L. M. L. Villani, “WAx: An integrated conceptual framework for the analysis of cyber-socio-technical systems,” *Saf Sci*, vol. 136, p. 105142, Apr. 2021.

[48] K. Siu *et al.*, “Architectural and Behavioral Analysis for Cyber Security,” *AIAA/IEEE Digital Avionics Systems Conference - Proceedings*, vol. 2019-September, Sep. 2019, doi: 10.1109/DASC43569.2019.9081652.

[49] Y. Brezhniev, “Multilevel Fuzzy Logic-Based Approach for Critical Energy Infrastructure’s Cyber Resilience Assessment,” *Conference Proceedings of 2019 10th International Conference on Dependable Systems, Services and Technologies, DESSERT 2019*, pp. 213–217, Jun. 2019, doi: 10.1109/DESSERT.2019.8770034.

[50] H. Haggi, R. R. Nejad, M. Song, and W. Sun, “A Review of Smart Grid Restoration to Enhance Cyber-Physical System Resilience,” 2019. doi: 10.1109/ISGT-Asia.2019.8881730.

[51] A. Shaked, L. Tabansky, and Y. Reich, “Incorporating Systems Thinking into a Cyber Resilience Maturity Model,” *IEEE Engineering Management Review*, vol. 49, no. 2, pp. 110–115, Apr. 2021, doi: 10.1109/EMR.2020.3046533.

[52] A. Andreasson and N. Fallen, “External cybersecurity incident reporting for resilience,” *Lecture Notes in Business Information Processing*, vol. 330, pp. 3–17, 2018, doi: 10.1007/978-3-319-99951-7_1.

[53] ISO/IEC 27001:2013, “ISO - ISO/IEC 27001:2013 - Information technology — Security techniques — Information security management systems — Requirements,” 2013.

[54] ISO/IEC 27002:2022, “ISO - ISO/IEC 27002:2022 - Information security, cybersecurity and privacy protection — Information security controls.” Accessed: Oct. 17, 2022. [Online]. Available: https://www.iso.org/standard/75652.html

[55] ISO/IEC 27031:2011, “ISO - ISO/IEC 27031:2011 - Information technology — Security techniques — Guidelines for information and communication technology readiness for business continuity.” Accessed: Oct. 17, 2022. [Online]. Available: https://www.iso.org/standard/44374.html

[56] ISO/IEC 27032:2012, “ISO - ISO/IEC 27032:2012 - Information technology — Security techniques — Guidelines for cybersecurity.” Accessed: Oct. 17, 2022. [Online]. Available: https://www.iso.org/standard/44375.html

[57] ISO/IEC 27701:2019, “ISO - ISO/IEC 27701:2019 - Security techniques — Extension to ISO/IEC 27001 and ISO/IEC 27002 for privacy information management — Requirements and guidelines.” Accessed: Oct. 17, 2022. [Online]. Available: https://www.iso.org/standard/71670.html

[58] Regulation (EU) GDPR, “General Data Protection Regulation (GDPR) – Official Legal Text,” 2016. Accessed: Oct. 17, 2022. [Online]. Available: https://gdpr-info.eu/

[59] NIS DIRECTIVE (EU), “DIRECTIVE (EU) 2016/ 1148 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL - of 6 July 2016 - concerning measures for a high common level of security of network and information systems across the Union,” 2016.

[60] National Institute of Standards, “NIST Cybersecurity Framework,” Gaithersburg, MD, Feb. 2014. Accessed: Nov. 18, 2020. [Online]. Available: https://doi.org/10.6028/NIST.CSWP.04162018

[61] NIST, “An Introduction to the Components of the Framework | NIST,” Apr. 16, 2018. https://www.nist.gov/cyberframework/online-learning/components-framework (accessed Nov. 24, 2022).

[62] CIS-Sapienza, “Framework Nazionale per la Cybersecurity e la Data Protection”, Accessed: Oct. 17, 2022. [Online]. Available: http://www.cybersecurityframework.it

[63] CIS-Sapienza, “Metodologia per il cybersecurity assessment con il Framework Nazionale per la Cybersecurity e la Data Protection”, Accessed: Oct. 17, 2022. [Online]. Available: http://www.cybersecurityfracamework.it

[64] K. D. Mitnick, W. L. Simon Foreword by Steve Wozniak, and F. Arynne, *The art of deception: Controlling the human element of security*. 2003. Accessed: Oct. 18, 2022. [Online]. Available: https://books.google.com/books?hl=en&lr=&id=rmvDDwAAQBAJ&oi=fnd&pg=PR7&ots=_fw-Px05U8&sig=-t9htwHF2coX5fdQjvfw82pd5Fw

[65] Verizon, “DBIR Data Breach Investigations Report,” 2022.

[66] IBM, “Cost of a data breach 2022 | IBM,” 2022. Accessed: Oct. 18, 2022. [Online]. Available: https://www.ibm.com/reports/data-breach

[67] CLUSIT, “Rapporto Clusit Marzo 2022 – Clusit,” 2022. Accessed: Oct. 15, 2022. [Online]. Available: https://clusit.it/rapporto-clusit/

[68] R. Naidoo, “A multi-level influence model of COVID-19 themed cybercrime,” *https://doi.org/10.1080/0960085X.2020.1771222*, vol. 29, no. 3, pp. 306–321, May 2020, doi: 10.1080/0960085X.2020.1771222.

[69] T. Li, X. Wang, and Y. Ni, “Aligning social concerns with information system security: A fundamental ontology for social engineering,” *Inf Syst*, vol. 104, Feb. 2022, doi: 10.1016/J.IS.2020.101699.

[70] F. Mouton, L. Leenen, and H. S. Venter, “Social engineering attack examples, templates and scenarios,” *Comput Secur*, vol. 59, pp. 186–209, Jun. 2016, doi: 10.1016/J.COSE.2016.03.004.

[71] Z. Wang, H. Zhu, P. Liu, and L. Sun, “Social engineering in cybersecurity: a domain ontology and knowledge graph application examples,” *Cybersecurity*, vol. 4, no. 1, pp. 1–21, Dec. 2021, doi: 10.1186/S42400-021-00094-6/FIGURES/27.

[72] D. M. Licht, D. J. Polzella, K. R. Boff, and H. G. Armstrong, “Human Factors, Ergonomics, and Human Factors Engineering: An Analysis of Definitions”.

[73] Z. Wang, H. Zhu, and L. Sun, “Social engineering in cybersecurity: Effect mechanisms, human vulnerabilities and attack methods,” *IEEE Access*, vol. 9, pp. 11895–11910, 2021, doi: 10.1109/ACCESS.2021.3051633.

[74] P. T. Costa and R. R. Mccrae, “The Five-Factor Model, Five-Factor Theory, and Interpersonal Psychology,” *Handbook of Interpersonal Psychology: Theory, Research, Assessment, and Therapeutic Interventions*, pp. 91–104, Mar. 2012, doi: 10.1002/9781118001868.CH6.

[75] P. Tetri and J. Vuorinen, “Dissecting social engineering,” *http://dx.doi.org/10.1080/0144929X.2013.763860*, vol. 32, no. 10, pp. 1014–1023, Oct. 2013, doi: 10.1080/0144929X.2013.763860.

[76] S. Uebelacker and S. Quiel, “The Social Engineering Personality Framework,” *Proceedings - 4th Workshop on Socio-Technical Aspects in Security and Trust, STAST 2014 - Co-located with 27th IEEE Computer Security Foundations Symposium, CSF 2014 in the Vienna Summer of Logic 2014*, pp. 24–30, Dec. 2014, doi: 10.1109/STAST.2014.12.

[77] W. Graterol, J. Diaz-Amado, Y. Cardinale, I. Dongo, E. Lopes-Silva, and C. Santos-Libarino, “Emotion detection for social robots based on nlp transformers and an emotion ontology,” *Sensors (Switzerland)*, vol. 21, no. 4, pp. 1–19, 2021, doi: 10.3390/s21041322.

[78] Y. Han and M. Moghaddam, “Analysis of sentiment expressions for user-centered design,” *Expert Syst Appl*, vol. 171, 2021, doi: 10.1016/j.eswa.2021.114604.

[79] K. M. Kwayu, V. Kwigizile, K. Lee, and J. S. Oh, “Discovering latent themes in traffic fatal crash narratives using text mining analytics and network topology,” *Accid Anal Prev*, vol. 150, Feb. 2021, doi: 10.1016/j.aap.2020.105899.

[80] X. Li, P. Penmetsa, J. Liu, A. Hainen, and S. Nambisan, “Severity of emergency natural gas distribution pipeline incidents: Application of an integrated spatio-temporal approach fused with text mining,” *J Loss Prev Process Ind*, vol. 69, Mar. 2021, doi: 10.1016/j.jlp.2020.104383.

[81] H. Abu Rasheed, C. Weber, J. Zenkert, P. Czerner, R. Krumm, and M. Fathi, “A Text Extraction-Based Smart Knowledge Graph Composition for Integrating Lessons Learned During the Microchip Design,” in *Advances in Intelligent Systems and Computing*, 2021, vol. 1251 AISC, pp. 594–610. doi: 10.1007/978-3-030-55187-2_43.

[82] J. Wang and C. C. Hsu, “A topic-based patent analytics approach for exploring technological trends in smart manufacturing,” *Journal of Manufacturing Technology Management*, vol. 32, no. 1, pp. 110–135, Sep. 2020, doi: 10.1108/JMTM-03-2020-0106.

[83] L. Barnewold and B. G. Lottermoser, “Identification of digital technologies and digitalisation trends in the mining industry,” *Int J Min Sci Technol*, vol. 30, no. 6, pp. 747–757, Nov. 2020, doi: 10.1016/j.ijmst.2020.07.003.

[84] F. Chiarello, N. Melluso, A. Bonaccorsi, and G. Fantoni, “A text mining based map of engineering design: Topics and their trajectories over time,” in *Proceedings of the International Conference on Engineering Design, ICED*, 2019, vol. 2019-August, pp. 2765–2774. doi: 10.1017/dsi.2019.283.

[85] T. Weißer, T. Saßmannshausen, D. Ohrndorf, P. Burggräf, and J. Wagner, “A clustering approach for topic filtering within systematic literature reviews,” *MethodsX*, vol. 7, 2020, doi: 10.1016/j.mex.2020.100831.

[86] B. Hari Prasath, S. Karthikeyan, and G. Mary Valantina, “Computerized highway defects recognition and classification system,” *International Journal of Pharmacy and Technology*, vol. 8, no. 1, pp. 11038–11048, 2016.

[87] T. Thinsungnoen, N. Kaoungku, P. Durongdumronchai, K. Kerdprasop, and N. Kerdprasop, "The Clustering Validity with Silhouette and Sum of Squared Errors," 2015, pp. 44–51. doi: 10.12792/iciae2015.012.

[88] The International Ergonomics Association, "What Is Ergonomics (HFE)?" https://iea.cc/what-is-ergonomics/ (accessed Nov. 06, 2022).

[89] S. A. Shappell and D. A. Wiegmann, *A Human Error Approach to Aviation Accident Analysis: The Human Factors Analysis and Classification System*. 2012. Accessed: Nov. 06, 2022. [Online]. Available: https://books.google.it/books/about/A_Human_Error_Approach_to_Aviation_Accid.html?id=28B3o2RRY9sC&redir_esc=y

[90] G. Dupont, "Avoid the dirty dozen with safety nets," *AIRBEATMAGAZINE*, 2009.

[91] D. N. Poller, M. Bongiovanni, B. Cochand-Priollet, S. J. Johnson, and M. Perez-Machado, "A human factor event-based learning assessment tool for assessment of errors and diagnostic accuracy in histopathology and cytopathology," *J Clin Pathol*, vol. 73, no. 10, pp. 681–685, Oct. 2020, doi: 10.1136/JCLINPATH-2020-206538.

[92] A. G. A. Samad, M. K. Johari, and S. Omar, "Preventing human error at an approved training organization using Dirty Dozen," *International Journal of Engineering and Technology(UAE)*, vol. 7, no. 4, pp. 71–73, 2018, doi: 10.14419/IJET.V7I4.13.21332.

[93] G. Desolda, L. S. Ferro, A. Marrella, T. Catarci, and M. F. Costabile, "Human Factors in Phishing Attacks: A Systematic Literature Review," *ACM Comput Surv*, vol. 54, no. 8, Nov. 2022, doi: 10.1145/3469886.

[94] A. Oltramari, D. Henshel, M. Cains, and B. Hoffman, "Towards a human factors ontology for cyber security," *CEUR Workshop Proc*, vol. 1523, pp. 26–33, 2015.

[95] M. Gratian, S. Bandi, M. Cukier, J. Dykstra, and A. Ginther, "Correlating human traits and cyber security behavior intentions," *Comput Secur*, vol. 73, pp. 345–358, Mar. 2018, doi: 10.1016/J.COSE.2017.11.015.

[96] Chartered Institute of Ergonomics & Human Factors, "The role of human factors in delivering cyber security," 2022. Accessed: Nov. 06, 2022. [Online]. Available: https://ergonomics.org.uk/resource/the-role-of-human-factors-in-delivering-cyber-security.html

[97] E. J. Williams, J. Hinds, and A. N. Joinson, "Exploring susceptibility to phishing in the workplace," *International Journal of Human Computer Studies*, vol. 120, pp. 1–13, Dec. 2018, doi: 10.1016/J.IJHCS.2018.06.004.

[98] S. Uebelacker and S. Quiel, "The social engineering personality framework," *Proceedings - 4th Workshop on Socio-Technical Aspects in Security and Trust, STAST 2014 - Co-located with 27th IEEE Computer Security Foundations Symposium, CSF 2014 in the Vienna Summer of Logic 2014*, pp. 24–30, Dec. 2014, doi: 10.1109/STAST.2014.12.

[99] L. Hadlington, "Human factors in cybersecurity; examining the link between Internet addiction, impulsivity, attitudes towards cybersecurity, and risky cybersecurity behaviours," *Heliyon*, vol. 3, no. 7, Jul. 2017, doi: 10.1016/J.HELIYON.2017.E00346.

[100] Science of Security and Privacy, “Science of Security Annual Report 2015 | CPS-VO,” 2015. Accessed: Nov. 06, 2022. [Online]. Available: https://cps-vo.org/sos/annualreport2015

[101] M. Evans, L. A. Maglaras, Y. He, and H. Janicke, “Human behaviour as an aspect of cybersecurity assurance,” *Security and Communication Networks*, vol. 9, no. 17, pp. 4667–4679, Nov. 2016, doi: 10.1002/SEC.1657.

[102] L. S. Ferro, A. Marrella, and T. Catarci, “A Human Factor Approach to Threat Modeling,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12788 LNCS, pp. 139–157, 2021, doi: 10.1007/978-3-030-77392-2_10/FIGURES/3.

[103] H. Young, T. van Vliet, J. van de Ven, S. Jol, and C. Broekman, “Understanding human factors in cyber security as a dynamic system,” *Advances in Intelligent Systems and Computing*, vol. 593, pp. 244–254, 2018, doi: 10.1007/978-3-319-60585-2_23.

[104] Chartered Institute of Ergonomics & Human Factors, “Human Affected Cyber Security Framework,” 2022. Accessed: Nov. 06, 2022. [Online]. Available: https://ergonomics.org.uk/resource/human-affected-cyber-security-framework.html

[105] D. Embrey, “Introduction to PIF’s Performance Influencing Factors (PIFs),” 2000.

[106] K. M. Groth and A. Mosleh, “A data-informed PIF hierarchy for model-based human reliability analysis,” *Reliab Eng Syst Saf*, vol. 108, pp. 154–174, Dec. 2012, doi: 10.1016/J.RESS.2012.08.006.

[107] K. Parsons, A. McCormac, M. Butavicius, M. Pattinson, and C. Jerram, “Determining employee awareness using the Human Aspects of Information Security Questionnaire (HAIS-Q),” *Comput Secur*, vol. 42, pp. 165–176, 2014, doi: 10.1016/J.COSE.2013.12.003.

[108] A. Pollini *et al.*, “Leveraging human factors in cybersecurity: an integrated methodological approach,” *Cognition, Technology and Work*, vol. 24, no. 2, pp. 371–390, May 2022, doi: 10.1007/S10111-021-00683-Y.

[109] P. K. Yeng, M. A. Fauzi, and B. Yang, “A Comprehensive Assessment of Human Factors in Cyber Security Compliance toward Enhancing the Security Practice of Healthcare Staff in Paperless Hospitals,” *Information (Switzerland)*, vol. 13, no. 7, Jul. 2022, doi: 10.3390/INFO13070335.

[110] E. Albrechtsen and J. Hovden, “Improving information security awareness and behaviour through dialogue, participation and collective reflection. An intervention study,” *Comput Secur*, vol. 29, no. 4, pp. 432–445, Jun. 2010, doi: 10.1016/J.COSE.2009.12.005.

[111] B. Khan, K. Alghathbar, S. I. Nabi, and M. Khan, “Effectiveness of information security awareness methods based on psychological theories,” *undefined*, vol. 5, no. 26, Oct. 2011, doi: 10.5897/AJBM11.067.

[112] K. Parsons, A. McCormac, M. Butavicius, M. Pattinson, and C. Jerram, “Determining employee awareness using the Human Aspects of Information Security Questionnaire (HAIS-Q),” *Comput Secur*, vol. 42, pp. 165–176, 2014, doi: 10.1016/J.COSE.2013.12.003.

[113] H. A. Kruger and W. D. Kearney, “A prototype for assessing information security awareness,” *Comput Secur*, vol. 25, no. 4, pp. 289–296, Jun. 2006, doi: 10.1016/J.COSE.2006.02.008.

[114] K. J. Knapp, R. Franklin Morris, T. E. Marshall, and T. A. Byrd, “Information security policy: An organizational-level process model,” *Comput Secur*, vol. 28, no. 7, pp. 493–508, Oct. 2009, doi: 10.1016/J.COSE.2009.07.001.

[115] R. Ait, M. Lahcen, B. Caulkins, R. Mohapatra, and M. Kumar, “Cybersecurity Review and insight on the behavioral aspects of cybersecurity”, doi: 10.1186/s42400-020-00050-w.

[116] S. Dekker and E. Hollnagel, “Human factors and folk models,” *Cognition, Technology & Work 2003 6:2*, vol. 6, no. 2, pp. 79–86, Oct. 2003, doi: 10.1007/S10111-003-0136-9.

[117] V. Zimmermann and K. Renaud, “Moving from a ‘human-as-problem” to a ‘human-as-solution” cybersecurity mindset,” *International Journal of Human Computer Studies*, vol. 131, pp. 169–187, Nov. 2019, doi: 10.1016/j.ijhcs.2019.05.005.

[118] PWC, “Key findings from The Global State of Information Security® Survey 2018,” 2018. Accessed: Nov. 07, 2022. [Online]. Available: www.pwc.com/gsiss

[119] NIST, “Success Stories | NIST,” Apr. 16, 2018. https://www.nist.gov/cyberframework/success-stories (accessed Nov. 07, 2022).

[120] L. E. Potter and G. Vickers, “What skills do you need to work in cyber security? A look at the Australian market,” *SIGMIS-CPR 2015 - Proceedings of the 2015 ACM SIGMIS Conference on Computers and People Research*, pp. 67–72, Jun. 2015, doi: 10.1145/2751957.2751967.

[121] C. Röcker, “Informal Communication and Awareness in Virtual Teams Why We Need Smart Technologies to Support Distributed Teamwork,” *Communications in Information Science and Management Engineering (CISME)*, 2012.

[122] G. A. Yukl, *Leadership in Organizations*, 8th Editio. Prentice-Hall, Upper Saddle River., 2013.

[123] A. Tkalac Verčič, D. Verčič, and K. Sriramesh, “Internal communication: Definition, parameters, and the future,” *Public Relat Rev*, vol. 38, no. 2, pp. 223–230, Jun. 2012, doi: 10.1016/J.PUBREV.2011.12.019.

[124] R. Smaliukienė and A. Survilas, “Relationship between organizational communication and creativity: How it advances in rigid structures?,” *Creativity Studies*, vol. 11, no. 1, pp. 230–243, Mar. 2018, doi: 10.3846/CS.2018.4004.

[125] L. L. Lemon, “The employee experience: how employees make meaning of employee engagement,” *Journal of Public Relations Research*, vol. 31, no. 5–6, pp. 176–199, Nov. 2019, doi: 10.1080/1062726X.2019.1704288.

[126] P. M. Wright and G. C. Mcmahan, “P-HRM: The Combination of PM and HRM,” *J Manage*, vol. 18, no. 2, pp. 295–320, Jul. 2016, doi: 10.1177/014920639201800205.

[127] S. M. Kompaso and M. S. Sridevi, “Employee Engagement: The Key to Improving Performance,” *International Journal of Business and Management*, vol. 5, no. 12, Nov. 2010, doi: 10.5539/IJBM.V5N12P89.

[128] A. Reeves, P. Delfabbro, and D. Calic, “Encouraging Employee Engagement With Cybersecurity: How to Tackle Cyber Fatigue,” *Sage Open*, vol. 11, no. 1, 2021, doi: 10.1177/21582440211000049.

[129] A. Rahim, “Managing Conflict in Organizations, Third Edition,” *Quorum Books*, 2001. https://books.google.it/books/about/Managing_Conflict_in_Organizations_Third.html?id=0HjZAQAACAAJ&redir_esc=y (accessed Nov. 07, 2022).

[130] Y. Bao, F. Zhu, Y. Hu, and N. Cui, “The Research of Interpersonal Conflict and Solution Strategies,” *Psychology*, vol. 07, no. 04, pp. 541–545, 2016, doi: 10.4236/PSYCH.2016.74055.

[131] I. Ellefsen, “The development of a cyber security policy in developing regions and the impact on stakeholders,” *2014 IST-Africa Conference and Exhibition, IST-Africa 2014*, 2014, doi: 10.1109/ISTAFRICA.2014.6880605.

[132] C. Stedman, “The Ultimate Guide to Cybersecurity Planning for Businesses,” 2022. Accessed: Nov. 07, 2022. [Online]. Available: https://www.techtarget.com/searchsecurity/The-ultimate-guide-to-cybersecurity-planning-for-businesses

[133] M. Pattinson, M. Butavicius, K. Parsons, A. McCormac, D. Calic, and C. Jerram, “The information security awareness of bank employees,” *Proceedings of the 10th International Symposium on Human Aspects of Information Security and Assurance, HAISA 2016*, pp. 189–198, 2016.

[134] S. Frey, “How to Eliminate the Prevailing Ignorance and Complacency Around Cybersecurity,” *Cybersecurity Best Practices*, pp. 1–10, 2018, doi: 10.1007/978-3-658-21655-9_1.

[135] J. Kawall, *On Complacency*. American Philosophical Quarterly, 2006. Accessed: Nov. 07, 2022. [Online]. Available: https://www.jstor.org/stable/20010257

[136] D. Henshel, M. G. Cains, B. Hoffman, and T. Kelley, “Trust as a Human Factor in Holistic Cyber Security Risk Assessment,” *Procedia Manuf*, vol. 3, pp. 1117–1124, 2015, doi: 10.1016/J.PROMFG.2015.07.186.

[137] A. A. Cain, “Trust and complacency in cyber security,” *Computer Science*, Jun. 2016, doi: 10.31979/ETD.A4NF-Q57U.

[138] T. F. Stafford, “Platform-Dependent Computer Security Complacency: The Unrecognized Insider Threat,” *IEEE Trans Eng Manag*, 2021, doi: 10.1109/TEM.2021.3058344.

[139] M. P. Sallos, A. Garcia-Perez, D. Bedford, and B. Orlando, “Strategy and organisational cybersecurity: a knowledge-problem perspective,” *Journal of Intellectual Capital*, vol. 20, no. 4, pp. 581–597, Oct. 2019, doi: 10.1108/JIC-03-2019-0041.

[140] M. A. Sasse, S. Brostoff, and D. Weirich, “Transforming the ‘weakest link’ - A human/computer interaction approach to usable and effective security,” *BT Technology Journal*, vol. 19, no. 3, pp. 122–131, Jul. 2001, doi: 10.1023/A:1011902718709.

[141] P. Menard, G. J. Bott, and R. E. Crossler, “User Motivations in Protecting Information Security: Protection Motivation Theory Versus Self-Determination Theory,” *Journal of Management Information Systems*, vol. 34, no. 4, pp. 1203–1230, Oct. 2017, doi: 10.1080/07421222.2017.1394083.

[142] S. M. Jaigirdar, S. Das, A. R. Chowdhury, S. Ahmed, and R. K. Chakrabortty, “Multi-objective multi-echelon distribution planning for perishable goods supply chain: a case study,” *https://doi.org/10.1080/23302674.2021.2020367*, 2022, doi: 10.1080/23302674.2021.2020367.

[143] I. Aouadni and A. Rebai, “Decision support system based on genetic algorithm and multi-criteria satisfaction analysis (MUSA) method for measuring job satisfaction,” *Ann Oper Res*, vol. 256, no. 1, pp. 3–20, Sep. 2017, doi: 10.1007/S10479-016-2154-Z/FIGURES/7.

[144] M. Pinzone *et al.*, “A framework for operative and social sustainability functionalities in Human-Centric Cyber-Physical Production Systems,” *Comput Ind Eng*, vol. 139, p. 105132, Jan. 2020.

[145] A. McIlwraith, “Information security and employee behaviour: How to reduce risk through employee education, training and awareness,” *Information Security and Employee Behaviour: How to Reduce Risk Through Employee Education, Training and Awareness*, pp. 1–195, Aug. 2021, doi: 10.4324/9780429281785.

[146] K. Øien, S. Massaiu, R. K. Tinmannsvik, and F. Størseth, “Development of early warning indicators based on Resilience Engineering,” *10th International Conference on Probabilistic Safety Assessment and Management 2010, PSAM 2010*, vol. 3, pp. 1762–1771, 2010.

[147] B. Bulgurcu, H. Cavusoglu, and I. Benbasat, “Information security policy compliance: An empirical study of rationality-based beliefs and information security awareness,” *MIS Q*, vol. 34, no. SPEC. ISSUE 3, pp. 523–548, 2010, doi: 10.2307/25750690.

[148] M. T. Siponen, “Conceptual foundation for organizational information security awareness,” *Information Management and Computer Security*, vol. 8, no. 1, pp. 31–41, 2000, doi: 10.1108/09685220010371394.

[149] Microsoft, “Microsoft Security Intelligence Report Volume 23 Supplement Malware at Microsoft 2,” 2018.

[150] J. C. Olivares-Rojas, E. Reyes-Archundia, J. A. Gutiérrez-Gnecchi, I. Molina-Moreno, J. Cerda-Jacobo, and A. Méndez-Patiño, “A methodology for cyber hygiene in smart grids,” *Dyna (Spain)*, vol. 97, no. 1, pp. 92–97, Jan. 2022, doi: 10.6036/10085.

[151] N. Nachin, C. Tangmanee, and K. Piromsopa, “How to Increase Cybersecurity Awareness,” *ISACA JOURNAL*, 2016, Accessed: Nov. 07, 2022. [Online]. Available: https://www.isaca.org/resources/isaca-journal/issues/2019/volume-2/how-to-increase-cybersecurity-awareness

[152] L. Razmerita, K. Kirchner, and P. Nielsen, “What factors influence knowledge sharing in organizations? A social dilemma perspective of social media communication,” *Journal of Knowledge Management*, vol. 20, no. 6, pp. 1225–1246, 2016, doi: 10.1108/JKM-03-2016-0112.

[153] R. L. Burkhead, B. J. Sharum, and S. A. Brown, “A PHENOMENOLOGICAL STUDY OF INFORMATION SECURITY INCIDENTS EXPERIENCED BY INFORMATION SECURITY PROFESSIONALS PROVIDING CORPORATE INFORMATION SECURITY INCIDENT MANAGEMENT,” 2014.

[154] N. H. Chowdhury, M. T. P. Adam, and G. Skinner, “The impact of time pressure on cybersecurity behaviour: a systematic literature review,” *Behaviour and Information Technology*, vol. 38, no. 12, pp. 1290–1308, Dec. 2019, doi: 10.1080/0144929X.2019.1583769.

[155] European Union Agency for Networked and Information Security, “ENISA Threat Landscape Report 2018 — ENISA,” 2018. https://www.enisa.europa.eu/publications/enisa-threat-landscape-report-2018 (accessed Jan. 29, 2021).

[156] C. D. Goldin, “Human Capital,” *Handbook of Cliometrics, ed. Claude Diebolt and Michael Haupert, 55-86. Heidelberg, Germany: Springer Verlag*, 2016, Accessed: Nov. 07, 2022. [Online]. Available: https://dash.harvard.edu/handle/1/34309590

[157] L. P. Tan and K. Y. Wong, “Linkage between knowledge management and manufacturing performance: a structural equation modeling approach,” *Journal of Knowledge Management*, vol. 19, no. 4, pp. 814–835, Jul. 2015, doi: 10.1108/JKM-11-2014-0487.

[158] L. Miarmi and K. G. DeBono, “The impact of distractions on heuristic processing: Internet advertisements and stereotype use,” *J Appl Soc Psychol*, vol. 37, no. 3, pp. 539–548, Mar. 2007, doi: 10.1111/J.1559-1816.2007.00173.X.

[159] P. Caponnetto, L. Bergefurt, M. Weijs-Perrée, C. Maris, and R. Appel-Meulenbroek, “Analyzing the Effects of Distractions While Working from Home on Burnout Complaints and Stress Levels among Office Workers during the COVID-19 Pandemic,” *Medical Sciences Forum 2021, Vol. 4, Page 44*, vol. 4, no. 1, p. 44, Jan. 2021, doi: 10.3390/ECERPH-3-09075.

[160] R. J. Simonson, J. R. Keebler, M. Lessmiller, T. Richards, and J. C. Lee, “Cybersecurity Teamwork: A Review of Current Practices and Suggested Improvements,” *https://doi.org/10.1177/1071181320641101*, vol. 64, no. 1, pp. 451–455, Feb. 2021, doi: 10.1177/1071181320641101.

[161] J. R. Gaufin, K. I. Kennedy, and E. D. Struthers, “Practical and affordable ways to cultivate leadership in your organization,” *Journal of Public Health Management and Practice*, vol. 16, no. 2, pp. 156–161, Mar. 2010, doi: 10.1097/PHH.0B013E3181C8CB63.

[162] K. M. Rogers and B. E. Ashforth, “Respect in Organizations: Feeling Valued as ‘We’ and ‘Me,’” *J Manage*, vol. 43, no. 5, pp. 1578–1608, May 2017, doi: 10.1177/0149206314557159.

[163] M. Bassanino, T. Fernando, and K. C. Wu, “Can virtual workspaces enhance team communication and collaboration in design review meetings?,” *http://dx.doi.org/10.1080/17452007.2013.775102*, vol. 10, no. 3–4, pp. 200–217, 2014, doi: 10.1080/17452007.2013.775102.

[164] M. A. Cavanaugh, W. R. Boswell, M. v. Roehling, and J. W. Boudreau, “An empirical examination of self-reported work stress among U.S. managers,” *Journal of Applied Psychology*, vol. 85, no. 1, pp. 65–74, 2000, doi: 10.1037/0021-9010.85.1.65.

[165] P. R. Scholtes, “The new competencies of leadership,” *Total Quality Management*, vol. 10, no. 4–5, pp. 704–710, 1999, doi: 10.1080/0954412997721.

[166] A. Ghadge, M. Weiß, N. D. N. D. Caldwell, and R. Wilding, *Managing cyber risk in supply chains: a review and research agenda*, vol. 25, no. 2. Emerald Group Publishing Ltd., 2019, pp. 223–240. doi: 10.1108/SCM-10-2018-0357.

[167] H. Taherdoost, “Understanding Cybersecurity Frameworks and Information Security Standards—A Review and Comprehensive Overview,” *Electronics (Switzerland)*, vol. 11, no. 14, Jul. 2022, doi: 10.3390/ELECTRONICS11142181.

[168] A. Zarreh, H. da Wan, Y. Lee, C. Saygin, and R. al Janahi, “Cybersecurity concerns for total productive maintenance in smart manufacturing systems,” *Procedia Manuf*, vol. 38, pp. 532–539, 2019, doi: 10.1016/J.PROMFG.2020.01.067.

[169] D. Beaty, *The Naked Pilot: The Human Factor in Aircraft Accidents*. 2011. Accessed: Nov. 07, 2022. [Online]. Available: https://books.google.it/books?hl=en&lr=&id=66R8AwAAQBAJ&oi=fnd&pg=PT8&dq=The+Naked+Pilot&ots=0gp9dkp6Mj&sig=VRSEqbeKnBmYMfdeUWS_q9bMlrA&redir_esc=y#v=onepage&q=The Naked Pilot&f=false

[170] A. Surprises, N. Sarter, D. Woods, and C. Billings, “Automation Surprises,” *Joint Cognitive Systems*, pp. 113–142, Mar. 2001, doi: 10.1201/9781420005684.CH10.

[171] A. Alahmari and B. Duncan, “Cybersecurity Risk Management in Small and Medium-Sized Enterprises: A Systematic Review of Recent Evidence,” *2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment, Cyber SA 2020*, Jun. 2020, doi: 10.1109/CYBERSA49311.2020.9139638.

[172] N. P. Podsakoff, J. A. Lepine, and M. A. Lepine, “Differential challenge stressor-hindrance stressor relationships with job attitudes, turnover intentions, turnover, and withdrawal behavior: A meta-analysis,” *Journal of Applied Psychology*, vol. 92, no. 2, pp. 438–454, Mar. 2007, doi: 10.1037/0021-9010.92.2.438.

[173] E. Mosley and S. Laborde, “Performing under Pressure: Influence of Personality-Trait-Like Individual Differences,” *Performance Psychology: Perception, Action, Cognition, and Emotion*, pp. 291–314, 2016, doi: 10.1016/B978-0-12-803377-7.00018-1.

[174] J. D’Arcy, A. Hovav, and D. Galletta, “User awareness of security countermeasures and its impact on information systems misuse: A deterrence approach,” *Information Systems Research*, vol. 20, no. 1, pp. 79–98, 2009, doi: 10.1287/isre.1070.0160.

[175] J. C. Giger and G. Pochwatko, “Sometimes it is not so bad to decide in a hurry: Infuence of different levels of temporal opportunity on the elaboration of purchasing intention,” *Polish Psychological Bulletin*, vol. 39, no. 4, pp. 209–216, Jan. 2008, doi: 10.2478/V10059-008-0026-3.

[176] J. Wang, T. Herath, R. Chen, A. Vishwanath, and H. R. Rao, “Research article phishing susceptibility: An investigation into the processing of a targeted spear phishing email,” *IEEE Trans Prof Commun*, vol. 55, no. 4, pp. 345–362, 2012, doi: 10.1109/TPC.2012.2208392.

[177] N. Nthala and I. Flechais, “‘If it’s urgent or it is stopping me from doing something, then i might just go straight at it’: A study into home data security decisions,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10292 LNCS, pp. 123–142, 2017, doi: 10.1007/978-3-319-58460-7_9.

[178] S. A. Woods and J. A. Sofat, “Personality and engagement at work: The mediating role of psychological meaningfulness,” *J Appl Soc Psychol*, vol. 43, no. 11, pp. 2203–2210, Nov. 2013, doi: 10.1111/JASP.12171.

[179] O. K. Osariemen, M. D. Chiahemba, S. S. Anshir, and W. Samuel, “The Role of Assertive Skills on Employees Work Performance: as key to Organisational Sustainabality,” 2021.

[180] S. J. Johnson, “Assertive Community Treatment: Evidence-Based Practice or Managed Recovery?,” *Assertive Community Treatment*, pp. 1–14, Sep. 2018, doi: 10.4324/9781351297486-1/ASSERTIVE-COMMUNITY-TREATMENT-EVIDENCE-BASED-PRACTICE-MANAGED-RECOVERY-SANDRA-JOHNSON.

[181] M. Weiss, “Effects of work stress and social support on information systems managers*,” *MIS Quarterly*, vol. 7, no. 1, pp. 29–43, Mar. 1983, doi: 10.2307/249075.

[182] H. C. Pham, “Information security burnout: Identification of sources and mitigating factors from security demands and resources,” *Journal of Information Security and Applications*, vol. 46, pp. 96–107, Jun. 2019, doi: 10.1016/J.JISA.2019.03.012.

[183] C. Nobles, “Stress, Burnout, and Security Fatigue in Cybersecurity: A Human Factors Problem,” *HOLISTICA – Journal of Business and Public Administration*, vol. 13, no. 1, pp. 49–72, Jul. 2022, doi: 10.2478/HJBPA-2022-0003.

[184] C. Brod, “Managing technostress: optimizing the use of computer technology.,” *Pers J*, vol. 61, no. 10, pp. 753–757, Oct. 1982.

[185] Cambridge Dictionary, “AWARENESS | English meaning - Cambridge Dictionary,” 2022. https://dictionary.cambridge.org/dictionary/english/awareness (accessed Nov. 11, 2022).

[186] L. S. Ferro, A. Marrella, T. Catarci, F. Sapio, A. Parenti, and M. de Santis, “AWATO: A Serious Game to Improve Cybersecurity Awareness,” pp. 508–529, 2022, doi: 10.1007/978-3-031-05637-6_33.

[187] G. el Hajal, R. Abi Zeid Daou, and Y. Ducq, “Human Firewall: Cyber Awareness using WhatApp AI Chatbot,” *2021 IEEE 3rd International Multidisciplinary Conference on Engineering Technology, IMCET 2021*, pp. 66–70, 2021, doi: 10.1109/IMCET53404.2021.9665642.

[188] Y. C. Fung and L. K. Lee, “A Chatbot for Promoting Cybersecurity Awareness,” *Lecture Notes in Networks and Systems*, vol. 370, pp. 379–387, 2022, doi: 10.1007/978-981-16-8664-1_33.

[189] SANS, “2022 Security Awareness Report | SANS Security Awareness,” 2022. Accessed: Nov. 08, 2022. [Online]. Available: https://go.sans.org/lp-wp-2022-sans-security-awareness-report

[190] L. O. Mailloux, M. L. Span, R. F. Mills, and W. L. Young, “A top down approach for eliciting systems security requirements for a notional autonomous space system,” 2019. doi: 10.1109/SYSCON.2019.8836929.

[191] B. Uchendu, J. R. C. Nurse, M. Bada, and S. Furnell, “Developing a cyber security culture: Current practices and future needs,” *Comput Secur*, vol. 109, Oct. 2021, doi: 10.1016/J.COSE.2021.102387.

[192] F. L. Al-Dawod, B. Stefanska, and R. Yakob, “The importance of risk awareness in cybersecurity among companies : A perspective on the role of top management,” 2021, Accessed: Nov. 08, 2022. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-177218

[193] L. W. Wong, V. H. Lee, G. W. H. Tan, K. B. Ooi, and A. Sohal, “The role of cybersecurity and policy awareness in shifting employee compliance attitudes: Building supply chain capabilities,” *Int J Inf Manage*, vol. 66, Oct. 2022, doi: 10.1016/J.IJINFOMGT.2022.102520.

[194] S. Boyson, “Cyber supply chain risk management: Revolutionizing the strategic control of critical IT systems,” *Technovation*, vol. 34, no. 7, pp. 342–353, 2014, doi: 10.1016/J.TECHNOVATION.2014.02.001.

[195] R. Smith, “Building A Cyber Supply Chain Assurance Reference Model A collaborative research project between SAIC and the,” *undefined*, 2009.

[196] J. K. Deane, C. L. Rees, and W. H. Baker, “Assessing the information technology security risk in medical supply chains,” *International Journal of Electronic Marketing and Retailing*, vol. 3, no. 2, pp. 145–155, 2010, doi: 10.1504/IJEMR.2010.032871.

[197] R. Koppel, J. Blythe, V. Kothari, and S. Smith, “Beliefs about Cybersecurity Rules and Passwords: A Comparison of Two Survey Samples of Cybersecurity Professionals Versus Regular Users.” 2016.

[198] I. Flechais and M. A. Sasse, “Stakeholder involvement, motivation, responsibility, communication: How to design usable security in e-Science,” *International Journal of Human Computer Studies*, vol. 67, no. 4, pp. 281–296, Apr. 2009, doi: 10.1016/J.IJHCS.2007.10.002.

[199] S. M. Furnell, A. Jusoh, and D. Katsabas, “The challenges of understanding and using security: A survey of end-users,” *Comput Secur*, vol. 25, no. 1, pp. 27–35, 2006, doi: 10.1016/J.COSE.2005.12.004.

[200] H. Heath and S. Cowley, “Developing a grounded theory approach: A comparison of Glaser and Strauss,” *Int J Nurs Stud*, vol. 41, no. 2, pp. 141–150, 2004, doi: 10.1016/S0020-7489(03)00113-5.

[201] K. H. Guo, Y. Yuan, N. P. Archer, and C. E. Connelly, “Understanding nonmalicious security violations in the workplace: A composite behavior model,” *Journal of Management Information Systems*, vol. 28, no. 2, pp. 203–236, Oct. 2011, doi: 10.2753/MIS0742-1222280208.

[202] NCSC, “Secure communications principles - NCSC.GOV.UK,” in *Strategic Marketing Management*, 2021. Accessed: Nov. 08, 2022. [Online]. Available: https://www.ncsc.gov.uk/guidance/secure-communication-principles

[203] R. Azmi, W. Tibben, and K. T. Win, “Review of cybersecurity frameworks: context and shared concepts,” *https://doi.org/10.1080/23738871.2018.1520271*, vol. 3, no. 2, pp. 258–283, May 2018, doi: 10.1080/23738871.2018.1520271.

[204] P. Rajivan, M. Champion, N. J. Cooke, S. Jariwala, G. Dube, and V. Buchanan, “Effects of teamwork versus group work on signal detection in cyber defense teams,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8027 LNAI, pp. 172–180, 2013, doi: 10.1007/978-3-642-39454-6_18/COVER.

[205] W. U. Hassan *et al.*, “NODOZE: Combatting Threat Alert Fatigue with Automated Provenance Triage”, doi: 10.14722/ndss.2019.23349.

[206] G. N. Angafor, I. Yevseyeva, and Y. He, “Game-based learning: A review of tabletop exercises for cybersecurity incident response training,” *Security and Privacy*, vol. 3, no. 6, p. e126, Nov. 2020, doi: 10.1002/SPY2.126.

[207] S. v. Veneruso, L. S. Ferro, A. Marrella, M. Mecella, and T. Catarci, “CyberVR: An Interactive Learning Experience in Virtual Reality for Cybersecurity Related Issues,” in *ACM International Conference Proceeding Series*, Sep. 2020. doi: 10.1145/3399715.3399860.

[208] Y. Cheng, Y. Sagduyu, J. Deng, J. Li, and P. Liu, “Integrated situational awareness for cyber attack detection, analysis, and mitigation,” in *Proceedings Volume 8385, Sensors and Systems for Space Applications V; 83850N (2012)*, May 2012, vol. 8385, pp. 169–179. doi: 10.1117/12.919261.

[209] SANS, “Improving Incident Response Through Simplified Lessons Learned Data Capture | SANS Institute,” 2021. Accessed: Nov. 08, 2022. [Online]. Available: https://www.sans.org/white-papers/40145/?utm_medium=Print&utm_source=SANS EDU Newsletter&utm_campaign=Research Review Journal

[210] NERC North American Electric Reliability Corporation, “Reputation Management: Corporate Image and Communication,” 2020.

[211] M. Bartock, J. Cichonski, M. Souppaya, M. Smith, G. Witte, and K. Scarfone, “Guide for Cybersecurity Event Recovery,” Dec. 2016, doi: 10.6028/NIST.SP.800-184.

[212] M. Angela Sasse, “‘Technology should be smarter than this!’: A vision for overcoming the great authentication fatigue,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 8425 LNCS, pp. 33–36, 2014, doi: 10.1007/978-3-319-06811-4_7.

[213] J. W. Hobbs, R. L. Burguete, P. F. Heyes, and E. A. Patterson, “Effect of eccentric loading on the fatigue performance of high-tensile bolts,” *Int J Fatigue*, vol. 22, no. 6, pp. 531–538, 2000, doi: 10.1016/S0142-1123(00)00004-9.

[214] R. S. Dalal, D. J. Howard, R. J. Bennett, C. Posey, S. J. Zaccaro, and B. J. Brummel, “Organizational science and cybersecurity: abundant opportunities for research at the interface,” *J Bus Psychol*, vol. 37, no. 1, Feb. 2022, doi: 10.1007/S10869-021-09732-9.

[215] N. Poehlmann, K. M. Caramancion, I. Tatar, Y. Li, M. Barati, and T. Merz, “The Organizational Cybersecurity Success Factors: An Exhaustive Literature Review,” in *Transactions on Computational Science and Computational Intelligence book series (TRACOSCI)*, Springer, Cham, 2021, pp. 377–395. doi: 10.1007/978-3-030-71017-0_27.

[216] K. Huang and K. Pearlson, “For what technology can’t fix: Building a model of organizational cybersecurity culture,” *Proceedings of the Annual Hawaii International Conference on System Sciences*, vol. 2019-Janua, pp. 6398–6407, 2019, doi: 10.24251/hicss.2019.769.

[217] A. Shah, R. Ganesan, S. Jajodia, and C. A. M. Hasan, “An outsourcing model for alert analysis in a cybersecurity operations center,” *ACM Transactions on the Web*, vol. 14, no. 1, Jan. 2020, doi: 10.1145/3372498.

[218] J. Saleem, B. Adebisi, R. Ande, and M. Hammoudeh, “A state of the art survey - Impact of cyber attacks on SME’s,” in *ACM International Conference Proceeding Series*, Jul. 2017, vol. Part F130522. doi: 10.1145/3102304.3109812.

[219] S. Armenia, M. Angelini, F. Nonino, G. Palombi, and M. F. Schlitzer, “A dynamic simulation approach to support the evaluation of cyber risks and security investments in SMEs,” *Decis Support Syst*, vol. 147, 2021, doi: 10.1016/j.dss.2021.113580.

[220] Panemon, “2018 State of Cybersecurity in Small & Medium Size Businesses,” 2018.

[221] J. Doran, G. Ryan, J. Bourke, and F. Crowley, “In-house or outsourcing skills: How best to manage for innovation?,” *International Journal of Innovation Management*, vol. 24, no. 1, Jan. 2020, doi: 10.1142/S1363919620500103.

[222] C. Ikerionwu, D. Edgar, and E. Gray, “The development of service provider’s BPO-IT framework,” *Business Process Management Journal*, vol. 23, no. 5, pp. 897–917, 2017, doi: 10.1108/BPMJ-10-2015-0146.

[223] C. C. Wolverton, R. Hirschheim, W. C. Black, and J. Burleson, “Outsourcing success in the eye of the beholder: Examining the impact of expectation confirmation theory on IT outsourcing,” *Information & Management*, vol. 57, no. 6, p. 103236, Sep. 2020, doi: 10.1016/J.IM.2019.103236.

[224] M. Benaroch, “Cybersecurity Risk in IT Outsourcing—Challenges and Emerging Realities,” pp. 313–334, 2020, doi: 10.1007/978-3-030-45819-5_13.

[225] E. Kweon, H. Lee, S. Chai, and K. Yoo, “The Utility of Information Security Training and Education on Cybersecurity Incidents: An empirical evidence,” *Information Systems Frontiers*, vol. 23, no. 2, pp. 361–373, 2021, doi: 10.1007/s10796-019-09977-z.

[226] C.-W. Liu, P. Huang, and H. C. Lucas, “Centralized IT Decision Making and Cybersecurity Breaches: Evidence from U.S. Higher Education Institutions,” *Journal of Management Information Systems*, vol. 37, no. 3, pp. 758–787, 2020, doi: 10.1080/07421222.2020.1790190.

[227] R. Leszczyna, “Review of cybersecurity assessment methods: Applicability perspective,” *Comput Secur*, vol. 108, 2021, doi: 10.1016/j.cose.2021.102376.

[228] N. Tissir, S. el Kafhali, and N. Aboutabit, “Cybersecurity management in cloud computing: semantic literature review and conceptual framework proposal,” *Journal of Reliable Intelligent Environments 2020 7:2*, vol. 7, no. 2, pp. 69–84, Oct. 2020, doi: 10.1007/S40860-020-00115-0.

[229] C.-W. Liu, P. Huang, and H. C. Lucas, “IT Centralization, Security Outsourcing, and Cybersecurity Breaches: Evidence from the U.S. Higher Education,” in *ICIS 2017: Transforming Society with Digital Innovation*, 2018.

[230] H. Cavusoglu, B. Mishra, and S. Raghunathan, “The effect of internet security breach announcements on market value: Capital market reactions for breached firms and internet security developers,” *International Journal of Electronic Commerce*, vol. 9, no. 1, pp. 70–104, 2004, doi: 10.1080/10864415.2004.11044320.

[231] DingDerui, HanQing-Long, XiangYang, GeXiaohua, and ZhangXian-Ming, “A survey on security control and attack detection for industrial cyber-physical systems,” *Neurocomputing*, vol. 275, pp. 1674–1683, Jan. 2018, doi: 10.1016/J.NEUCOM.2017.10.009.

[232] S. M. Sabri, R. Sulaiman, A. Ahmad, and A. Tang, “A review on IT outsourcing practices for e-business transformation among SMEs in Malaysia,” *Conference Proceedings - 6th International Conference on Information Technology and Multimedia at UNITEN: Cultivating Creativity and Enabling Technology Through the Internet of Things, ICIMU 2014*, pp. 124–129, Mar. 2015, doi: 10.1109/ICIMU.2014.7066616.

[233] DibbernJens, GolesTim, HirschheimRudy, and JayatilakaBandula, “Information systems outsourcing,” *ACM SIGMIS Database: the DATABASE for Advances in Information Systems*, vol. 35, no. 4, pp. 6–102, Nov. 2004, doi: 10.1145/1035233.1035236.

[234] P. A. Laplante, T. Costello, P. Singh, S. Bindiganavile, and M. Landon, “The Who, What, Why, Where, and When of IT Outsourcing,” *IT Prof*, vol. 6, no. 1, pp. 19–23, Jan. 2004, doi: 10.1109/MITP.2004.1265538.

[235] K. G. 1966- Liao and L. A. 1967- Reátegui, “Information technology outsourcing in emerging markets,” 2002, Accessed: Jul. 27, 2021. [Online]. Available: https://dspace.mit.edu/handle/1721.1/26892

[236] H. Hussin, M. R. Hasan, and N. A. Molok, “Exploring the perception and practices of IT outsourcing among Malaysia SMEs: Receivers perspective,” *Proceeding of the 3rd International Conference on Information and Communication Technology for the Moslem World: ICT Connecting Cultures, ICT4M 2010*, 2010, doi: 10.1109/ICT4M.2010.5971928.

[237] T. Kern, “The Gestalt of an information technology outsourcing relationship: an exploratory analysis,” *Proceedings of the eighteenth international conference on Information systems*, vol. 3, pp. 37–58, 1997, doi: 10.5555/353071.353085.

[238] M. Almutairi and S. Riddle, “State of the art of IT outsourcing and future needs for managing its security risks,” *2018 International Conference on Information Management and Processing, ICIMP 2018*, vol. 2018-Janua, pp. 42–48, Mar. 2018, doi: 10.1109/ICIMP1.2018.8325839.

[239] M. M. Rajaeian, A. Cater-Steel, and M. Lane, “A systematic literature review and critical assessment of model-driven decision support for IT outsourcing,” *Decis Support Syst*, vol. 102, pp. 42–56, Oct. 2017, doi: 10.1016/J.DSS.2017.07.002.

[240] S. Dhar, “From outsourcing to Cloud computing: evolution of IT services,” *Management Research Review*, vol. 35, no. 8, pp. 664–675, Jul. 2012, doi: 10.1108/01409171211247677.

[241] B. A. Aubert and S. Rivard, “The Outsourcing of IT Governance,” pp. 43–59, 2020, doi: 10.1007/978-3-030-45819-5_3.

[242] H. Liang, J. J. Wang, Y. Xue, and X. Cui, “IT outsourcing research from 1992 to 2013: A literature review based on main path analysis,” *Information & Management*, vol. 53, no. 2, pp. 227–251, Mar. 2016, doi: 10.1016/J.IM.2015.10.001.

[243] J. Barthélemy and D. Geyer, “An empirical investigation of IT outsourcing versus quasi-outsourcing in France and Germany,” *Information & Management*, vol. 42, no. 4, pp. 533–542, May 2005, doi: 10.1016/J.IM.2004.02.005.

[244] H. U. Rahman, M. Raza, P. Afsar, M. Khan, N. Iqbal, and H. U. Khan, “Making the Sourcing Decision of Software Maintenance and Information Technology,” *IEEE Access*, vol. 9, pp. 11492–11510, 2021, doi: 10.1109/ACCESS.2021.3051023.

[245] M. R. S. Reyes, “A systematic review of the literature on information technology outsourcing services,” *J Phys Conf Ser*, vol. 1513, no. 1, p. 012007, Mar. 2020, doi: 10.1088/1742-6596/1513/1/012007.

[246] T. Tsygankova, O. Yatsenko, O. Mozgovyy, T. Didukh, and L. Patsola, “Mobilization of innovative and resource factors for d evelopment of national outsourcing IT companies,”

Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu, no. 1, pp. 191–197, 2021, doi: 10.33271/NVNGU/2021-1/191.

[247] J. M. Marco-Simó and J. A. Pastor-Collado, “IT Outsourcing in the Public Sector: A Descriptive Framework from a Literature Review,” *https://doi.org/10.1080/1097198X.2019.1701357*, vol. 23, no. 1, pp. 25–52, Jan. 2019, doi: 10.1080/1097198X.2019.1701357.

[248] H. Bauer, G. Sherf, and V. von der Tann, “Six ways CEOs can promote cybersecurity in the IoT age | McKinsey,” 2017. https://www.mckinsey.com/featured-insights/internet-of-things/our-insights/six-ways-ceos-can-promote-cybersecurity-in-the-iot-age (accessed Jul. 27, 2021).

[249] M. Lezzi, M. Lazoi, and A. Corallo, “Cybersecurity for Industry 4.0 in the current literature: A reference framework,” *Comput Ind*, vol. 103, pp. 97–110, Dec. 2018, doi: 10.1016/J.COMPIND.2018.09.004.

[250] E. Amir, S. Levi, and T. Livne, “Do firms underreport information on cyber-attacks? Evidence from capital markets,” 2018, doi: 10.1007/s11142-018-9452-4.

[251] S. Nahavandi, “Industry 5.0-a human-centric solution,” *Sustainability (Switzerland)*, vol. 11, no. 16, Aug. 2019, doi: 10.3390/SU11164371.

[252] E. Commission, D.-G. for R. and Innovation, M. Breque, L. de Nul, and A. Petridis, *Industry 5.0 : towards a sustainable, human-centric and resilient European industry*. Publications Office, 2021. doi: doi/10.2777/308407.

[253] S. Wellsandta, Z. Rusak, S. Ruiz Arenas, D. Aschenbrenner, K. A. Hribernik, and K.-D. Thoben, “Concept of a Voice-Enabled Digital Assistant for Predictive Maintenance in Manufacturing,” *SSRN Electronic Journal*, Oct. 2020, doi: 10.2139/SSRN.3718008.

[254] Z. Zhu *et al.*, “AR-mentor: Augmented reality based mentoring system,” *ISMAR 2014 - IEEE International Symposium on Mixed and Augmented Reality - Science and Technology 2014, Proceedings*, pp. 17–22, Nov. 2014, doi: 10.1109/ISMAR.2014.6948404.

[255] J. J. Roldán, E. Crespo, A. Martín-Barrio, E. Peña-Tapia, and A. Barrientos, “A training system for Industry 4.0 operators in complex assemblies based on virtual reality and process mining,” *Robot Comput Integr Manuf*, vol. 59, pp. 305–316, Oct. 2019, doi: 10.1016/J.RCIM.2019.05.004.

[256] F. Costantino, A. Falegnami, L. Fedele, M. Bernabei, S. Stabile, and R. Bentivenga, “New and Emerging Hazards for Health and Safety within Digitalized Manufacturing Systems,” *Sustainability 2021, Vol. 13, Page 10948*, vol. 13, no. 19, p. 10948, Oct. 2021, doi: 10.3390/SU131910948.

[257] M. Casillo, F. Colace, L. Fabbri, M. Lombardi, A. Romano, and D. Santaniello, “Chatbot in industry 4.0: An approach for training new employees,” *Proceedings of 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, TALE 2020*, pp. 371–376, Dec. 2020, doi: 10.1109/TALE48869.2020.9368339.

[258] D. Bohus and A. I. Rudnicky, “LARRI: A Language-Based Maintenance and Repair Assistant,” pp. 203–218, 2005, doi: 10.1007/1-4020-3075-4_12.

[259] C. Li, J. Park, H. Kim, and D. Chrysostomou, “How can i help you? An intelligent virtual assistant for industrial robots,” *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 220–224, 2021, doi: 10.1145/3434074.3447163.

[260] C. Li and H. J. Yang, “Bot-X: An AI-based virtual assistant for intelligent manufacturing,” *Multiagent and grid system*, vol. 17, pp. 1–14, 2021.

[261] T. Y. Chen, Y. C. Chiu, N. Bi, and R. T. H. Tsai, “Multi-modal Chatbot in Intelligent Manufacturing,” *IEEE Access*, vol. 9, 2021, doi: 10.1109/ACCESS.2021.3083518.

[262] J. N. Pires, “Robot-by-voice: Experiments on commanding an industrial robot using the human voice,” *Industrial Robot*, vol. 32, no. 6, pp. 505–511, 2005, doi: 10.1108/01439910510629244/FULL/XML.

[263] N. Ade, N. Quddus, T. Parker, and S. C. Peres, “ProBot – A Procedure Chatbot for Digital Procedural Adherence:,” *https://doi.org/10.1177/1071181320641054*, vol. 64, no. 1, pp. 224–228, Feb. 2021, doi: 10.1177/1071181320641054.

[264] M. Zimmer, A. Al-Yacoub, P. Ferreira, and N. Lohse, “Towards Human-Chatbot Interaction: A Virtual Assistant for the Ramp-up Process,” *UKRAS20 Conference: “Robots into the real world” Proceedings*, vol. 3, pp. 108–110, May 2020, doi: 10.31256/QX5DT5V.

[265] F. Longo and A. Padovano, “Voice-enabled Assistants of the Operator 4.0 in the Social Smart Factory: Prospective role and challenges for an advanced human–machine interaction,” *Manuf Lett*, vol. 26, pp. 12–16, Oct. 2020, doi: 10.1016/J.MFGLET.2020.09.001.

[266] M. E. Foster and C. Matheson, “Following Assembly Plans in Cooperative, Task-Based Human-Robot Dialogue,” in *Conference: Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue (Londial 2008)*, 2008.

[267] S. Wellsandta, Z. Rusak, S. Ruiz Arenas, D. Aschenbrenner, K. A. Hribernik, and K.-D. Thoben, “Concept of a Voice-Enabled Digital Assistant for Predictive Maintenance in Manufacturing,” *SSRN Electronic Journal*, Oct. 2020, doi: 10.2139/SSRN.3718008.

[268] J. S. Jwo, C. S. Lin, and C. H. Lee, “An Interactive Dashboard Using a Virtual Assistant for Visualizing Smart Manufacturing,” *Mobile Information Systems*, vol. 2021, 2021, doi: 10.1155/2021/5578239.

[269] M. Casillo, F. Colace, L. Fabbri, M. Lombardi, A. Romano, and D. Santaniello, “Chatbot in industry 4.0: An approach for training new employees,” *Proceedings of 2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, TALE 2020*, pp. 371–376, Dec. 2020, doi: 10.1109/TALE48869.2020.9368339.

[270] A. Augello, G. Pilato, A. Machi, and S. Gaglio, “An approach to enhance chatbot semantic power and maintainability: Experiences within the FRASI project,” *Proceedings - IEEE 6th International Conference on Semantic Computing, ICSC 2012*, pp. 186–193, 2012, doi: 10.1109/ICSC.2012.26.

[271] S. Mantravadi, A. D. Jansson, and C. Møller, “User-Friendly MES Interfaces: Recommendations for an AI-Based Chatbot Assistance in Industry 4.0 Shop Floors,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 12034 LNAI, pp. 189–201, 2020, doi: 10.1007/978-3-030-42058-1_16/FIGURES/6.

[272] M. Kiruthiga Devi, M. S. Divakar, V. Vimal Kumar, M. D. E. Jaincy, R. A. Kalpana, and S. R. M. Kumar, “FARMER’S assistant using ai voice bot,” *2021 3rd International Conference on Signal*

Processing and Communication, ICPSC 2021, pp. 527–531, May 2021, doi: 10.1109/ICSPC51351.2021.9451760.

[273] V. Nayak, P. R Nayak N, Sampoorna, Aishwarya, and N. H. Sowmya, “Agroxpert - Farmer assistant,” *Global Transitions Proceedings*, vol. 2, no. 2, pp. 506–512, Nov. 2021, doi: 10.1016/J.GLTP.2021.08.016.

[274] F. C. T. Wu, O. N. J. Hong, A. J. C. Trappey, and C. v. Trappey, “VR-enabled chatbot system supporting transformer mass-customization services,” *Advances in Transdisciplinary Engineering*, vol. 12, pp. 291–300, Sep. 2020, doi: 10.3233/ATDE200088.

[275] P. Chandar *et al.*, “Leveraging Conversational Systems to Assists New Hires During Onboarding,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 10514 LNCS, pp. 381–391, 2017, doi: 10.1007/978-3-319-67684-5_23.

[276] A. Lòpez, J. Sànchez-Ferreres, J. Carmona, and L. Padrò, “From Process Models to Chatbots,” *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11483 LNCS, pp. 383–398, 2019, doi: 10.1007/978-3-030-21290-2_24/FIGURES/6.

[277] M. Locatelli, E. Seghezzi, L. Pellegrini, L. C. Tagliabue, and G. M. Di Giuda, “Exploring natural language processing in construction and integration with building information modeling: A scientometric analysis,” *Buildings*, vol. 11, no. 12, Dec. 2021, doi: 10.3390/buildings11120583.

[278] K. H. Sharif and S. Y. Ameen, “A Review of Security Awareness Approaches with Special Emphasis on Gamification,” *3rd International Conference on Advanced Science and Engineering, ICOASE 2020*, pp. 151–156, Dec. 2020, doi: 10.1109/ICOASE51841.2020.9436595.

[279] A. Corallo, M. Lazoi, M. Lezzi, and A. Luperto, “Cybersecurity awareness in the context of the Industrial Internet of Things: A systematic literature review,” *Comput Ind*, vol. 137, May 2022, doi: 10.1016/J.COMPIND.2022.103614.

[280] D. R. Boccardo, L. M. S. Bento, and F. H. Costa, “Towards a practical information security maturity evaluation method focused on people, process and technology,” *2021 IEEE International Workshop on Metrology for Industry 4.0 and IoT, MetroInd 4.0 and IoT 2021 - Proceedings*, pp. 721–726, Jun. 2021, doi: 10.1109/METROIND4.0IOT51437.2021.9488471.