International Journal of Information Security
https://doi.org/10.1007/s10207-023-00769-w

SPECIAL ISSUE PAPER

Cyber threat assessment and management for securing healthcare ecosystems using natural language processing

Stefano Silvestri[1] · Shareful Islam[2,3] · Dmitry Amelin[5] · Gabriele Weiler[5] · Spyridon Papastergiou[3,4] · Mario Ciampi[1]

Abstract

The healthcare sectors have constantly faced significant challenge due to the rapid rise of cyber threats. These threats can pose any potential risk within the system context and disrupt the critical healthcare service delivery. It is therefore necessary for the healthcare organisations to understand and tackle the threats to ensure overall security and resilience. However, threats are continuously evolved and there is large amount of unstructured security-related textual information is available. This makes the threat assessment and management task very challenging. There are a number of existing works that consider Machine Learning models for detection and prediction of cyber attack but they lack of focus on the Natural Language Processing (NLP) to extract the threat information from unstructured security-related text. To this end, this work proposes a novel method to assess and manage threats by adopting natural language processing. The proposed method has been tailored for the healthcare ecosystem and allows to identify and assess the possible threats within healthcare information infrastructure so that appropriate control and mitigation actions can be taken into consideration to tackle the threat. In detail, NLP techniques are used to extract the useful threat information related to specific assets of the healthcare ecosystems from the largely available security-related information on Internet (e.g. cyber security news), to evaluate the level of the identified threats and to select the required mitigation actions. We have performed experiments on real healthcare ecosystems in Fraunhofer Institute for Biomedical Engineering, considering in particular three different healthcare scenarios, namely implantable medical devices, wearables, and biobank, with the purpose of demonstrating the feasibility of our approach, which is able to provide a realistic manner to identify and assess the threats, evaluate the threat level and suggest the required mitigation actions.

Keywords Cyber threat assessment · Cyber threat mitigation · Healthcare information infrastructure · Natural language processing · Artificial intelligence

This paper is an extended, improved version of the paper: Shareeful Islam, Spyridon Papastergiou, Stefano Silvestri - "Cyber Threat Analysis Using Natural Language Processing for a Secure Healthcare System". In Proceedings of IEEE Symposium on Computers and Communications (ISCC) 2022, Rhodes, Greece, IEEE, 2022. DOI: 10.1109/ISCC55528.2022.9912768.

✉ Stefano Silvestri
stefano.silvestri@icar.cnr.it

Shareful Islam
shareeful.islam@aru.ac.uk

Dmitry Amelin
dmitry.amelin@ibmt.fraunhofer.de

Gabriele Weiler
gabriele.weiler@ibmt.fraunhofer.de

Spyridon Papastergiou
spyrospapastergiou@gmail.com

Mario Ciampi
mario.ciampi@icar.cnr.it

[1] Institute for High Performance Computing and Networking of the National Research Council of Italy, ICAR-CNR, Via Pietro Castellino 111, 80131 Naples, Italy

[2] School of Computing and Information, Science Anglia Ruskin University, Cambridge, UK

[3] Focal Point, Waterloo, Belgium

[4] Department of Informatics, University of Piraeus, Piraeus, Greece

[5] Fraunhofer Institute for Biomedical Engineering IBMT, Sulzbach, Germany

Published online: 27 October 2023

1 Introduction

The healthcare sector is now rapidly evolving with continuous adoption of new technologies including Internet of Things (IoT) and connected medical devices to support medical applications and healthcare service delivery. In particular, the Healthcare Information Infrastructure (HCII) is undergoing a significant technological revolution that not only offers advantages to the overall healthcare processes but also expands the attack surface for threat actors to exploit potential vulnerabilities.

The sector is constantly suffering with a number of successful cyber attacks in recent years, including NHS ransomware attacks in 2017 and the Ireland's Department of Health and Health Service Executive in 2021 [1]. Moreover, there are intrinsic vulnerabilities in the medical devices, such as flaws in Braun's infusion pump or Medtronic insulin pump, that could pose potential threat to the patient's health [2]. Hence, nearly 90% of healthcare organisations have experienced a data breach in 2018 [3]. Therefore, there is a pressing need to identify and analyse the cyber threats relevant within the healthcare ecosystem, so that appropriate controls action can be taken into consideration for secure and resilient healthcare service delivery [4].

However, understanding threats for a specific context is a challenging task, due to the evolving nature of the threat and availability of large amounts of unstructured Natural Language (NL) Cyber Security (CS) text in various sources such as on blog posts, CS news websites, social media, and others. This text often contains crucial information related to the assets of the HCII, which is difficult to extract threat-related information due to the complexity nature of the NL due to the presentation of polysemy, irony, complex, long sentences and nonstandard abbreviations or acronyms [5].

Some of the existing works contribute to address these challenges notably using Deep Learning (DL) for Natural Language Processing (NLP) to implement Named Entity Recognition (NER) that aims to identify and classifies the security-related name entities in the NL text [6–9]. But lack of focus on the analysing and prioritising the threats so that suitable mitigation actions are considered. It is also necessary to focus on the threats that are relevant to the healthcare sector. In this context, this paper presents an unique threat assessment and management approach for the Healthcare system based on the extraction of threat-related information from NL CS textual documents, using a BERT-based architecture [10] to implement a NER and a text classification modules. This work presents the outputs produced by the AI4HEALTHSEC EU project,[1] which aims to develop solutions to tackle the security and privacy threats in Healthcare ICT Infrastructures, and extends our initial idea presented in [11–13], as well as includes the results obtained in a pilot developed in real setting.

The paper makes three important contributions. Firstly, it focuses on the holistic understanding of the threats that potentially affect the healthcare ecosystem. The identified threats are analysed and prioritised so that suitable controls actions can be identified to tackle the threats. Secondly, the proposed approach adopts NLP techniques to extract possible threats that are related to the HCII assets from unstructured NL sources, determines the corresponding level of the identified threat, and if possible suggests the corresponding mitigation actions. Finally, it shows the results of the experiments performed in a real-world healthcare ecosystem scenario from Fraunhofer Institute for Biomedical Engineering (IBMT) with three assets, i.e. implantable medical devices, wearables, and biobank, assessing the applicability and the usefulness of the proposed work. The results show that the proposed approach determines and categorises possible threats based on the large collection of textual CS news that has been specifically linked with the IBMT assets.

The paper is structured as follows. In the next Sect. 2, the most recent related work presented in literature are summarised. Then, the details of the proposed approach are described in Sect. 3, which is followed by the description of the resources and datasets required to implement our method in Sect. 4. Then, Sect. 5 describes the healthcare ecosystem scenarios and the experimental assessment, showing and discussing the obtained results. Finally, Sect. 6 outlines the conclusions and the possible future works.

2 Related works

This section provides an overview of existing works which are relevant to our work. In particular, we examine the areas of cyber threats in the Healthcare sector, threat modelling and analysis, ML models for threat analysis, and NLP methods exploited for CS tasks.

2.1 Cyber threats in healthcare sector

Healthcare sectors are constantly facing challenge to tackle the evolving and sophisticated cyber threats. In fact, the healthcare domain is one of the most critical ones and requires specific approaches to prevent, identify and assess attacks. A recent study showed that at least 20% of the medical device manufacturers experienced ransomware or malware attacks in the last 20 months [14].

The cyber attacks have also targeted medical devices, such as infusion pumps, or healthcare services, such as medicine delivery of the healthcare system [15]. For instance, implantable cardiac devices get security features associated with the system architecture, which uses device-to-device

[1] https://www.ai4healthsec.eu.

authentication schemes such as hard-coded credentials on home monitoring devices for authenticating to patient support networks. An attacker can exploit this credential to access the network [16].

In addition to cyber attacks, which have been targeted against the vulnerabilities of information technology (IT) infrastructures, also social engineering-based attacks are breaching the security of Healthcare Organisations, with often severe outcomes [17]. Threat actors constantly target to obtain patient-sensitive information and hacking is considered one of the main causes that discloses patient-sensitive healthcare data [18]. Vulnerabilities in Medical IoT devices are now considered sources of threats and risks in the healthcare domain. Specifically, simulated attacks have been made on devices, including pacemakers, insulin pumps and drug infusion pumps [19]. Malware-related threats such as Medjack can inject malicious code into unprotected medical devices, which can impact other parts of the overall healthcare ICT infrastructure [20]. The Centre for Internet Security emphasises data breaches, DDoS, inside threats and business email compromise, and data breach as the key cyber attack in the healthcare sector [21]. The cyber attack path generation and analysis for securing the healthcare ecosystem is proposed by [22]. The work considers assets and their dependencies within the healthcare sectors and demonstrates how cyber attacks can be propagated from medical devices to other parts of the healthcare ecosystem.

2.2 Threat modelling and analysis

Threat modelling and analysis is a key activity to understand the possible threats that impact on the assets within the system context. There are a number of threat modelling approaches that are widely used for the threat modelling notably PASTA, STRIDE, Attack-tree, Threat Dragon.

The Attack Simulation and Threat Analysis (PASTA) a risk-centric approach for the identification of security flaws and possible impact, allowing to determine the more appropriate controls for the mitigation [23]. The model advocates analyst-business collaboration with the intent to assess, document, and propose countermeasures relative to the likelihood of an attack. STRIDE denotes spoofing, tampering, repudiation, information disclosure, denial of service (DoS), and elevation of privilege aims to aid reason, detect, and identify threats targeting a system by breaking down processes, data flows and stores, as well as trust boundaries [24]. Attack Tree follows a tree-based hierarchical structure to describe security of a system [25]. The root node considers the goal, while the lower level nodes consider the possible attack to the system. This tree provides potential attack patterns for specific targets, while describing threats aimed at a system and the possible counterattack approaches to realise them.

A data-driven threat analysis (d-TM) approach is proposed which focuses on analysing the threats from different abstractions of organisational data in three phases, including storage, process and transmit [26]. In particular, OWASP Threat Dragon aims to create threat model diagrams as visual indication of threat and possible attack surface related to the threat [27]. The threats are categorised by following STRIDE and CIA properties. It is an open-source tool that organisations of any type can use for threat modelling. Additionally, it adopts a rule engine to identify possible vulnerabilities within the infrastructure and visually analyse them through the tool. The flexibility of this tool allows it to consider all types of threat and supports determination of possible countermeasures to mitigate the threat and creating new or updating existing features to tackle the threats.

In [28], ESSecA (Expert System for Security Assessment) is presented which is a methodology that guides penetration testers during the assessment of IoT systems from the threat-intelligence perspective. The approach produces a Threat Model and a list of Attack Plans for each identified threat as a part of analysis. The use of Knowledge Bases (KBs) is crucial to support CS threat and risks modelling and assessment.

There are several works that highlight the threat modelling for healthcare system context. A report by Threatmodeller mentioned that the healthcare sector is primarily targeted for the threat actor not only for the data but also any connected medical devices, IoT and others [29]. The report highlighted the FDA guidance with six principles, i.e. cybersecurity is an integral part of device safety and the QSR, Security by design, Transparency, Security risk management, Security architecture, and Testing/objective evidence that need to be followed by the manufacturer to protect medical devices. A threat model is proposed tailored for the selected IoT health devices in [30] by combining STRIDE and DREAD model. In particular, threats are identified using STRIDE model based on the device access points and ranked using DREAD. The proposed model is applicable to all relevant stakeholders including designers and users of health IoT devices for enhancing overall security of the IoT health devices. The work considers a number of health devices notably Connected inhalers, Ingestible sensors, Leaf Healthcare ulcer sensor, Intelligent asthma monitoring system, etc and threats are identified and categorised based on STRIDE for the identified devices. A list of countermeasures are also identified to mitigate these threats.

The security and privacy challenges in Medical Cyber-Physical Systems (MCPS) due to unique application requirements and characteristics of MCPS and threat modelling of MCPS are discussed in [31]. The proposed trust and threat model considers MCPS Stakeholders, including healthcare practitioners, system administrator, and non-medical staff with levels of trust for these users such as Trustworthy users,

Trusted but error-prone, Untrustworthy, Temporarily trustworthy. Threats are considered based on communication links, software, platform and users perspectives. The work provides example architecture and possible mitigations relating to anomaly detection, cryptographic measures, system hardening of the identified threats.

2.3 Machine learning for cyber security

Machine Learning (ML) models are now widely considered by the research and industry community for cyber security analysis. In [32] a tool is proposed for detection of attacks targeting the Healthcare cyber-physical system devoted to patient health remote monitoring. The abnormal health features are detected using a multi-heuristic cyber ant optimisation-based feature extraction process and results showed the proposed approach effectively allows to monitor patient health conditions, detecting at the same time the data breaches, and improving cloud security. A cyber supply chain threat analysis using Random Forest and GBoost algorithms for the threat prediction is presented by [33]. The method considers threat intelligence and predicts the Tactics, Techniques, and Procedures (TTP) deployed for a cyber attack, leveraging Random Forest and XG Boost algorithms, providing a high accuracy in their experimental assessment.

SHChecker is another novel threat analysis framework recently proposed by [34], which combines ML and Formal Analysis capabilities for the Smart Healthcare Systems (SHSs). This framework focuses on Internet of Medical Things (IoMT) and adopts several ML algorithms, such as Decision Tree (DT), Artificial Neural Network (ANN), K-means, and others, in order to implement CS threat analysis. The obtained result showed that ANN-based algorithms provide less accuracy than DT-based algorithms. Another dimension of ML model is Deep Neural Networks (DNN) and Deep Learning which have been also successfully adopted for cyber security analysis.

The authors of [35] illustrated use of AI on teaching and training new algorithms for securing, preparing, and adapting the healthcare system. The main purpose is to develop a concept healthcare system supported by autonomous artificial intelligence that can use edge health devices with real-time data. The work also shows how Natural Language (NL) CS reports and other sources (social media, forums, etc.) can open a very interesting resource for the real-time analysis of CS-related data. Data and knowledge-driven CS Named Entity Recognition (NER) methodology is used as a DL Bidirectional Long Short Term Memory Conditional Random Field (BiLSTM-CRF) architecture by [9]. An improved approach to analyse the CS closed-domain texts is performed by [13], which integrates some KBs, to model the details of the assets (application, vendor, version, etc.) affected by CS issues. DL architecture is used for the identification of relevant CS information, such as vulnerabilities, attack discoveries and advanced persistent threats [6]. The proposed architecture is formed by a word-embedding layer, a BiLSTM-CRF layer, followed by an additional BiLSTM layer in output. The results of the experimental assessment demonstrated some improvements with respect to the baselines.

Another approach by [36] considers DL method for the analysis of the severity of CS threats which considers a corpus of 6000 tweets where software vulnerabilities are described, annotated with opinions towards their severity. Furthermore, it is also presented a method for linking software vulnerabilities reported in tweets to CVEs and NVD KBs. The results showed that a high precision in forecasting high-severity vulnerabilities, also highlighting that reports of severe vulnerabilities online are predictive of real-world exploits.

2.4 NLP for unstructured cyber security text analysis

The current state of the art NLP techniques exploit Deep Learning approaches based on the Transformer architecture [37], the BERT model [10] and its evolutions. These approaches have been recently also applied for the definition of CS NLP-based methodologies, allowing for the automatic analysis of NL documents. The authors of [38] presented a BERT model devoted to provide CS domain embeddings for analysing CS texts. They pretrained the model on a large corpus consisting of the text extracted from scientific papers, Twitter posts, CS-domain web pages, and vulnerability database, successfully testing it on several CS NLP tasks. Also in [39], a closed-domain CS BERT model, fine-tuned with a large corpus of textual CS data to recognise CS entities. The CyBERT model focuses on the identification of CS claims in Industrial Control Systems (ICS) NL documents [40, 41]. The authors collected a large corpus of labelled sequences from ICS device documentation to pretrain and fine-tune a BERT language model. The result shows that this model improves the results compared to other Transformer-based language models trained on generic domain corpora. The authors of [42] presented SecureBERT, a BERT model trained on a CS-domain large NL corpora, which proved to outperforms other similar models in NLP tasks in cybersecurity.

An open-source Python library for CS NER named CyNER is presented in [43]. This library has the purpose of leveraging the huge Open Cyber Threat Intelligence (OpenCTI) information from unstructured textual format available from several heterogeneous sources on the Internet. The proposed approach combines Transformer-based models for extracting CS entities, heuristics for extracting different indicators of compromise, and publicly available NER models for the extraction of generic entity types. A CS NER model based on a joint architecture which is com-

posed of a BERT-based model and an LSTM layer is used by [8] to extract character and text features and to predict sequence labels in NL text. This model is exploited to create a knowledge graph for Cyber Threat Intelligence (CTI) to rapidly analyse advanced cyber threats in threat situations from NL documents. The results showed that the proposed joint BERT models can significantly outperform the state-of-the-art methods.

A method for automatically structuring CTIs and converting them into standard STIX format is presented in [44]. In detail, the method exploits several NLP methodologies, such as Text Classification, NER, Relation Extraction, with the purpose of identifying Indicators Of Compromises (IOCs) within NL CTIs. The experiments demonstrated that it can extract IOCs that are not included in existing reputation sites, and that it can automatically extract IOCs that have been exploited for a long time and across multiple attack groups.

Another CS NER approach capable of obtaining state of the art results is described in [7]. The proposed model exploits a combined architecture that includes a BERT with Whole World Masking (BERT WWM) and a BiLSTM-CRF neural network. In [45] a dataset for Event Detection (ED) in CS texts called CASIE is presented, including both cyberattack and vulnerability-related events. The authors also implemented a methodology able to populate a semantic model, with the ultimate goal of integration into a knowledge graph of CS data. Moreover, they defined and trained different deep neural networks to perform the ED task, obtaining interesting results.

In summary, all these above-mentioned works contributed towards developing various methods and techniques for the threat analysis and recent attention towards adoption of ML, specifically NLP for the threat analysis, certainly made important advancement. However, there is a lack of focus on the healthcare sector and assessment and management of the extracted threats. To this end, our work contributes to extract threat-related information to the healthcare sector using CS NLP. Additionally, we also contribute to assess and manage the threats for ensuring security and resilience of the healthcare service delivery.

3 Proposed approach

The proposed unique threat assessment and approach aims to ensure security and resilience of the overall healthcare ecosystem. An important part of the overall method is to understand the healthcare context specifically Healthcare Information Infrastructure (HCII) and entities such as hospital, clinic and care home who are linked with each other to support the healthcare service delivery. It aims to assist healthcare institutions to understand the threats appropriate for their context, quantify the threats to determine the severity so that appropriate control actions can be taken into consideration to tackle the threats.

The approach investigates unstructured security-related data to extract threat-related information using Natural Language Processing (NLP), such data often contains critical information to understand the threats for a specific context. Additionally, we have also adopted widely used Common Attack Pattern Enumeration and Classification (CAPEC)[2] and Common Platform Enumeration (CPE).[3] The benefits of using these standards are that they provide a systematic way to identify and classify the threats and list the assets. The method considers three main phases, which are sequential with each other. It initiates from understanding healthcare ecosystem context so that threats within the context can be linked and mitigated. The phases are discussed below.

- Phase 1: Understand Healthcare Ecosystem Context
- Phase 2: Threat Identification and Assessment
- Phase 3: Threat Mitigation

3.1 Phase 1: Understand Healthcare Ecosystem Context

This first phase aims to understand the healthcare ecosystem so that assets and services of the ecosystem can be identified and analysed. This phase is required to determine the importance of each asset within the ecosystem and it is performed with the support of CS-domain Knowledge Bases. Healthcare ecosystem is a complex patient and doctor-based community system. The ecosystem is massive in size and consists of healthcare entities such as hospitals, care homes, or clinics and actors such as doctors, nurses and other practitioners that work together to deliver healthcare services.

The assets within the ecosystem are interconnected, for instance connected medical devices such as X-ray, infusion pump or insulin pump are connected with other infrastructure such as server and healthcare software [4]. Therefore healthcare ecosystem consists of a set of entities, such as hospitals, patients, practitioners, processes, and services that rely on the interconnected ICT infrastructure of four different areas of consideration, including: i) patient healthcare devices, ii) IT devices, iii) individual healthcare services and process, and iv) supply chain services.

This phase requires a close collaboration and active participation among the relevant stakeholders of the healthcare organisation for identification of assets and services. This involves engaging an IT service team (i.e. system administrator, infrastructure and security analyst) with healthcare practitioner, admin and business development team. The identification of assets and services involves conducting two

[2] https://capec.mitre.org.

[3] https://cpe.mitre.org/.

levels of interviews, i.e. strategic and the operational/ technical level. Strategic-level interviews target admin and business development team to identify healthcare services and processes with the system context, while operational/ technical level interviews target the operational team including system admin, and healthcare practitioners to understand the assets and other infrastructure to support the services The security analyst coordinates the activities along with other members of the IT service team. This phase includes two steps, which are briefly presented below.

3.1.1 Identify the healthcare services

The first step of the analysis involves the identification of the available healthcare services for an organisation. A comprehensive list of all healthcare services must be generated such as patient registration and appointment, operation schedule, blood text, X-ray,and many more. Service is viewed as a business process, where a collection of activities and tasks form a Business Flow, ensuring the proper operation of the service. Each business process is part of a specific healthcare ecosystem and may depend on external actors. Those dependencies must also be considered for the service identification. The services are necessary as they are linked with the assets of the overall system context. Hence, services are fully dependent on the assets for their delivery.

3.1.2 Identify and analyse the healthcare assets

Once the services are identified, then it is necessary to determine the possible assets which are linked with the service. The asset identification can also investigate the document related to medical devices such as European Database on Medical Devices (EUDAMED), which provides information about the medical devices for enhancing transparency and visualising the lifecycle of the specific device. Therefore, specific information about a device such as manufacture, conformity information, clinical studies can be obtained from EUDAMED and used for the asset which are linked with the identified services. Our approach advocates using the Common Platform Enumeration (CPE) Knowledge Base to map the identified assets with specific classes of applications, operating systems, and hardware devices. CPE provides a structure naming for the assets. The inventory tools and scanners can also assist to automatically identify the assets. Hence, CPE provides a structured naming scheme for all assets relevant for the healthcare ecosystem context. The identified assets are the internal system components that are controlled by the examined healthcare organisation(s). We have considered four distinct healthcare areas as presented in Table 1 to describe the assets within the HCII. Additionally, assets are also categorised depending on its functionalities, as shown in Table 2. This allows us to determine the importance of each asset within the ecosystem.

Table 1 Assets areas

Area	Name
1	User interactions with implants and sensors
2	Medical equipment and IT devices
3	Services and processes
4	Interdependent HCIIs - Ecosystem

Table 2 Assets categories

Category	Functionalities
Influence	Found in most organisations, distinct
Type	Software, hardware, Operating System (OS), information Sensitivity
Sensitivity	Restricted, unrestricted
Criticality	Essential, required, deferrable

Note that an asset may be involved in one or many services. For example, a patient may order a repeat prescription and at the same time request for an appointment which may link with multiple services simultaneously. Assets analysis determines the criticality of the assets based on their dependencies within the services. Hence, criticality is measured based on the dependency level that an asset has with other system components within these services. In general, four dependency levels are defined:

- *Independent assets* have a distinct operation and exhibit no dependency on other assets. If the asset fails, no cascading events occur.
- *Incoming dependency* of an asset, if syntactically another asset utilises its data or functionality. If such an asset fails, the operation of all related assets that utilise its data or functionality may be disrupted as well.
- *Outgoing dependency* of an asset, if syntactically it utilises data or functionality of another asset. Therefore, if the latter asset fails, the operation of the former asset will be affected as well.
- *Coupling relationship* reveals that two assets have both incoming and outgoing dependencies. Thereupon, failures in one of the assets will affect the functionality of the other.

Therefore, the more a specific asset is dependent, the higher is the criticality. An asset can have completely independent operation (independent dependency) or may only produce data (incoming dependency), only consume data (outcoming dependency), or produce and consume data for

a specific service (coupling dependency). We consider three distinct levels of criticality of the asset.

- *High*: the asset includes more outgoing and coupling dependencies with other asset and related service
- *Medium*: the asset includes less coupling but more incoming and outgoing dependency with other asset and related service
- *Low*: the asset includes no coupling dependency and less outgoing or incoming dependency.

3.2 Phase 2: Threat Identification and Assessment

Once all the assets are identified, phase 2 aims to identify and assess the possible threats related to the assets. Threats are the possible attempts to exploit the vulnerabilities, which may affect the identified healthcare services and assets. In particular, an individual or group of people known as threat actor attempting to gain access or exploit a vulnerability of healthcare asset or the damage caused to hinder the organisation's ability to provide its healthcare services. The threat actor includes various tactics and techniques to successfully execute the attack with a specific aim for the attack. The threats can be categorised through threat taxonomies and assessed in a qualitative manner using threat scales. This phase considers two main steps:

- Threat Identification;
- Threat Assessment.

3.2.1 Threat identification

This step identifies the threats for each asset of the HCII (as identified by the Healthcare Ecosystem Context Component). We consider threat intelligence data as knowledge base for this step.

Threat-related knowledge base There are several available sources that catalogue known threats along with their characteristics. We consider Common Attack Pattern Enumeration and Classification (CAPEC) to identify the threats relevant to the HCII and CAPEC can also link with assets using the Common Platform Enumeration (CPE) catalogue.A list of these considered characteristics is given below.

- *Abstraction*: Defines the different abstraction levels that apply to an attack pattern. A Meta level attack pattern provides an abstract characterisation of a specific methodology or technique used for an attack and generalisation of a related group of standard level attack pattern. It is often void of specific technology or implementation and provides an understanding of a high-level approach.
- *Status*: Defines the different status values of an entry of the CAPEC catalogue including view, category, attack pattern.
- *Description*: A short description of the threat.
- *Alternate Terms*: Indicates one or more other names used to describe this attack patterns.
- *Vendor and Item*: Respectively identify the vendor and item (e.g. *Google* and *Chrome*) affected by the CS issue.
- *Likelihood of Attack*: Determines the likelihood and severity of an attack that leverages using the attack pattern and may not be completely accurate for all attacks.
- *Typical Severity*: It is used to capture an overall average severity value for attacks that leverage this attack pattern with the understanding that it will not be completely accurate for all attacks.
- *Related Attack Patterns*: Refers to other attack patterns and related high-level categories. These relationships give insight to similar items that may exist at higher and lower levels of abstraction.
- *Execution Flow*: It is used to provide a detailed step-by-step flow performed by an adversary for a specific attack pattern. It is applicable to attack patterns with an abstraction level of details.
- *Prerequisites*: Indicates one or more prerequisite conditions necessary for an attack.
- *Skills and Resource Required*: Respectively describe skill level or knowledge and possible resources (e.g. CPU cycles, IP addresses, tools) required by an adversary for an attack.
- *Indicators*: The possible indicators including activities, events, conditions, or behaviours that may indicate an attack which could be imminent, in progress, or has occurred. Each Indicator element provides a textual description of the indicator.
- *Consequences*: The possible consequences associated with an attack pattern. The required Scope element identifies the security property that is violated. The optional Impact element describes the technical impact that arises if an adversary succeeds in their attack.
- *Mitigation*: The suitable counter measure to prevent or mitigate the risk of an attack. The approaches described in each mitigation element should help improve the resiliency of the target system, reduce its attack surface, or reduce the impact of the attack if it is successful.
- *Example Instances*: It is used to describe one or more example instances of the attack pattern. An example helps the reader understand the nature, context, and variability of the attack in more practical and concrete terms.
- *Related Weaknesses*: Contains references to weaknesses associated with this attack pattern. The association implies a weakness that must exist for a given attack to be successful. If multiple weaknesses are associated with the attack pattern, then any of the weaknesses (but

not necessarily all) may be present for the attack to be successful. Each related weakness is identified by a CWE identifier.
- *Taxonomy Mappings*: It is used to provide a mapping from an entry (Attack Pattern or Category) in CAPEC to an equivalent entry in a different taxonomy.
- *Notes*: It is used to provide any additional comments that cannot be captured using the other elements of the view.

Additionally, CAPEC provides mappings to threat classification taxonomies with the MITRE ATT&CK framework. Utilising the relationship between the two knowledge bases, the threats contained in CAPEC can be sorted using the tactics residing in ATT&CK, which express short-term adversary goals throughout the execution of an attack. The adoption of CPE and CAPEC in our approach provides a systematic way to identify the asset and link with the threat. Hence, the purpose of CPE is to provide a standardised format to enumerate the software and other assets, whereas CAPEC allows to understand the possible attack patterns based on domain and mechanisms of attack. Therefore, attack pattern targeting to specific products can be obtained by correlating among CPE and CAPEC. Note that, asset specifically medical devices may be not having any CPE entry so that cannot be correlated with attack pattern. But medical devices are increasingly reliant on technology for delivering healthcare services. Our approach considers four different types of dependency among the assets with four possible asset areas as presented in Table 1. Except independent type assets, all other categories focus on dependent assets. For instance, a medical device may connect with another IT device, platform, or communication infrastructure. Such dependency can pose any potential threat from other dependent assets. Therefore, even if the asset may not have any CPE entry or CAPEC entry, there is still a possibility for successful execution of any threat.

The outcome of the step is the list of identified threats for each asset that operate for the provision of each identified healthcare service. Each threat is listed along with a CAPEC ID, a CAPEC category that will be used to rate the threat, and a set of available characteristics that is both informative about the threat and able to procure material for extensions on the methodology throughout the second iteration. In summary, for each asset $a_{k,i}$ it was possible to identify all the corresponding threats $T_{i,a_{k,i}}$ in the context of a healthcare service.

3.2.2 Threat assessment

Once the threats are identified, it is necessary to assess them for determining corresponding severity levels. This step aims to assess the threats based on the natural language history of reported incidents related to those threats in an automated manner. Healthcare organisations can proactively determine the suitable controls to tackle the identified threats as explained in phase 3 based on the prioritised threats.

The history of reported incidents related to those threats is used for the threat prioritisation, namely the assignment of a level to each threat. We extract this history from the huge online natural language resources available on the Internet, such as forums, social media, news sites, and others. Therefore, we needed to implement a system for the automatic analysis of the various available NL sources. For this purpose, we exploited a Natural Language Processing (NLP) Named Entity Recognition (NER) AI architecture based on a BERT model [10], following the pipeline depicted in Fig. 1. In this schema, a set of input natural language sources corresponding for instance to threat reports, articles on various blogs/websites, Twitter data, online publicly available datasets, and/or log-files of the HCIIs, etc., can be fed as input into the NLP module, which extract the mentions of threats and assets from the input NL text, by means of a BERT-based NER module, specifically fine-tuned on this task.

Then, the level of the threats are calculated based on the occurrence of the mentions of that threat within the considered dataset. Finally, a mapping among the assets of the services of the HCIIs (identified by the previous Healthcare Ecosystem Context component) and the pairs asset/threat extracted through NLP with the corresponding threat level is performed, allowing in this way to leverage this information for threat prioritisation in the RA4Health methodology. In our experimental assessment we adopted a data set composed of natural language CS news, described in detail in the next Sect. 4.2.

In summary, the whole Threat Assessment task is performed by exploiting an AI-based NLP methodology, which consists of two main steps:

- *Extraction of threats*. This step is performed by leveraging a BERT-based NER model, specifically trained for the identification of potential threats and assets mentioned in natural language documents, mapping them with the assets and threats of the HCII previously identified.
- *Prioritising of threats*. The evaluation of the threat level asset/threat previously extracted through NLP NER, based on the percentage of occurrence in the considered NL documents.

In the first step, the threats and the corresponding assets are automatically extracted from NL texts by a NER module, which has been previously trained on CS domain-annotated documents. For this purpose, we adopted NER models for the CS domain based on Transformer architecture [37]. In particular, the results of previous experiments reported in

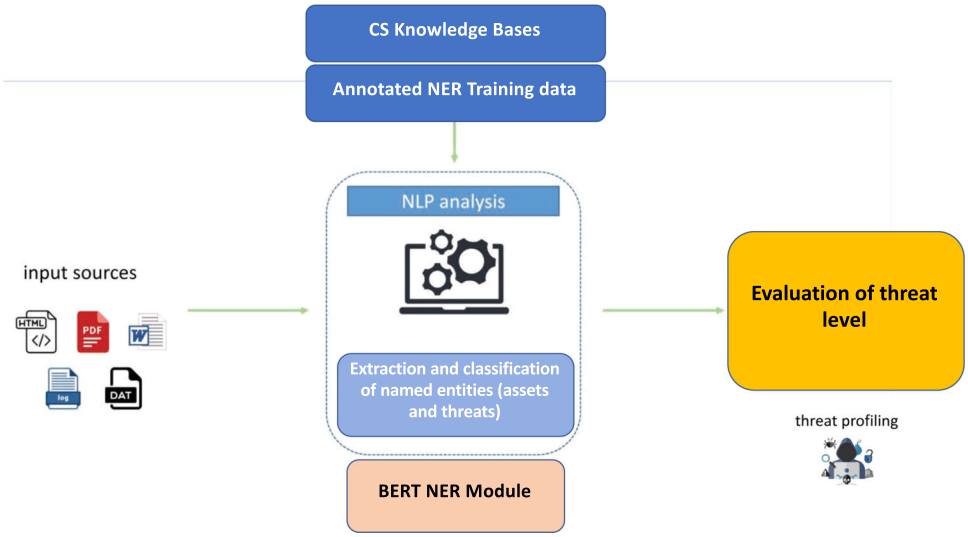

Fig. 1 Schema of the individual threat assessment component

[12, 13] suggested the use SecBERT[4] model, an NLM pre-trained on a large document collection belonging to the CS domain (as described more in details in the next Sect. 4.3). This NLM has been fine-tuned on a data set annotated with threats and assets. The annotation has been performed using a semisupervised iterative annotation method, based on neural network and KBs, described in [13, 46, 47].

After extracting the potential threats and the corresponding assets of the HCII from the CS natural language text, it is possible to evaluate the threat-level in the latter step of the proposed methodology. In this case, we correlate the level of a threat to the percentage of its occurrence in the large CS news documents we collected: the more frequently the threat is mentioned, the higher will be considered its threat level. We assigned five different threat levels, based on the percentage of occurrence in our dataset, as shown in Table 3, ranging from *Very High* to *Very Low*.

Although NLP ML-based approaches could provide some false positives and negatives (in our case, identifying false assets or threats entities or missing some of them in the analysed text), our method exploits the percentage of asset/threat pairs found within the considered corpora for threat prioritisation. Therefore, when a corpora is sufficiently large, a small number of false positives and negatives would not sensibly affect the final percentage calculated.

3.3 Phase 3: Threat mitigation

This final phase of our method aims to mitigate the identified threats based on the informed decision making taken into account the threats and their levels. This phase focuses on reviewing the existing controls and determining the additional controls necessary to mitigate the threats. As stated before, our approach considers the history of reported incidents related to the threats by following the online natural language resources available on the Internet, such as forums, social media, news sites, and others.

Note that the healthcare sector is heavily regulated by relevant legislations for protection of patient data, health and safety and medical devices, such as GDPR or HIPAA [48]. Due to the rapid evolution of the healthcare sector with the adoption of new technologies, cyber attacks are constantly increasing. Healthcare systems are now more regulated due to this technological evolution. Therefore, in case of any incident, the healthcare organisation needs to report the incident within a specific time frame in order to comply with the relevant legislation. Our approach focuses on the incidents which are reported and available. Furthermore, cyber security incidents from connected medical devices or other parts of the systems can be caused due to the potential vulnerabilities of specific products, which are published. When a healthcare organisation records an incident but is not reported, then the severity of such incident could be insignificant. Therefore, our approach focuses on the incidents which are reported and available.

Therefore, the control needs to be comprehensive to tackle all aspects of the prioritised threats. These controls are mainly security mechanisms with specific functionalities such as corrective, detective and preventive. Corrective controls mitigate the potential damage due to any successful threat. Preventative focuses to restrict the unwanted or unauthorised activities from occurring due to specific threat, where detective control identifies the possible anomaly within the overall system context which requires immediate action.

This phase determines the potential security measures by looking at the threats, their levels and related assets. However, the extraction of the range of occurrence can also support determining the root causes of the threats and com-

[4] https://github.com/jackaduma/SecBERT.

Table 3 Threat level and corresponding percentage of occurrence

Threat level	Occurrence percentage	Description
Very high	[80–100]	Severe impact on critical services and assets
High	[60–80]	Significant impact on critical services and assets
Medium	[40–60]	Intermediate impact on services and assets and no critical service would be affected
Low	[20–40]	Low impact and no critical service would be affected
Very low	[1–20]	Significant low impact

mon threat patterns. This allows to determine the appropriate control to tackle the threats and prioritise some controls. In particular, depending on threat level such as very high and high, identified controls may need immediate implementation. Controls should also be categorised based on their types such as technical, administrative, physical, compliance. The technical control uses hardware and software mechanisms as the basis for controlling the risks and associated vulnerabilities. Administrative control ensures policies, procedures, and standards that relate to personnel and business practice based on overall security needs. Physical control refers to a tangible entity that is used to prevent or detect unauthorised access to physical areas, systems, or assets. Finally, compliance control refers to the relevant controls, which are required to comply with relevant legislation. Our approach identifies and categorises the controls based on types and functionalities.

Figure 2 depicts the physical view and the pipeline of the proposed threat assessment and mitigation methodology, summarising the corresponding phases and components, showing the input and output data and the data flows, as well as the required resources. As shown in the Figure, the first phase of the methodology leverages the CPE KBs to obtain the assets list of the considered HCII, also providing the corresponding areas and category of each asset. Therefore, it is possible to identify the threats of each asset of the HCII, exploiting the CAPEC KB and the HCII asset list obtained in the previous phase. After that, the obtained assets and threats of the HCII have been identified and characterised. Then, it is possible to perform the threat assessment phase, exploiting a NER module (based on a BERT architecture) and the NL CS document corpus, which includes news from the CS domain extracted from the web. Finally, the assets and threats of the HCII and the corresponding threats level can be used to select the most suitable mitigation control.

In example, if an e-health medical instrument includes a set of subsystems (IoT devices, PC, software, etc.), the corresponding assets (databases, operating systems, web servers, etc.) are identified and categorised in the first phase of the proposed approach, exploiting the information available in the CPE KB. Then, it is also possible to identify the threats related to these assets, leveraging the CAPEC KBs. At this point, a threat assessment is obtained using textual natural language documents from the CS domain available on the web. In detail, using an NLP approach, it is possible to extract the information related to the assets and threats. The threat assessment is based on the percentage of mentions of the pair asset/threat found within these documents. Finally, the obtained information can be used to select the most appropriate mitigation actions, such as patching or upgrading a software, changing the configuration of an operating system, and others.

4 Resources and implementation

The implementation of the methodology described in previous Sect. 3 requires several resources and datasets. In particular, some knowledge bases are needed for the threat identification phase, and also for the annotation of the training set to fine-tune the NLM for the NER task. Specific datasets must be used to train the NLM, as well as to make available a natural language source to extract the information for the threat assessment phase. Moreover, it is necessary to use a fine-tuned NLM for the NER task. Finally, some other tools and resources are required, such as NLP preprocessing tools, specific scripts and software libraries. This Section describes the details of the resources collected and used to implement the proposed approach.

4.1 Knowledge bases

As already mentioned in previous Sect. 3.2.1, the CAPEC and CVE KBs have been leveraged in the threat identification phase, allowing to model assets and threats in the HCII, and creating a list of detected threats for each asset that operates for the provision of each identified healthcare service. Moreover, these two KBs were also used to support the annotation of the NER training set, by means of Distant Supervision (DS) and Active Learning (AL) iterative annotation process described in previous works [12, 13, 46, 47], which allows to annotate a dataset with few efforts, if an NLM and domain-specific KBs are available. In this way, it is possible to address the lack of annotated training set for the specific CS domain.

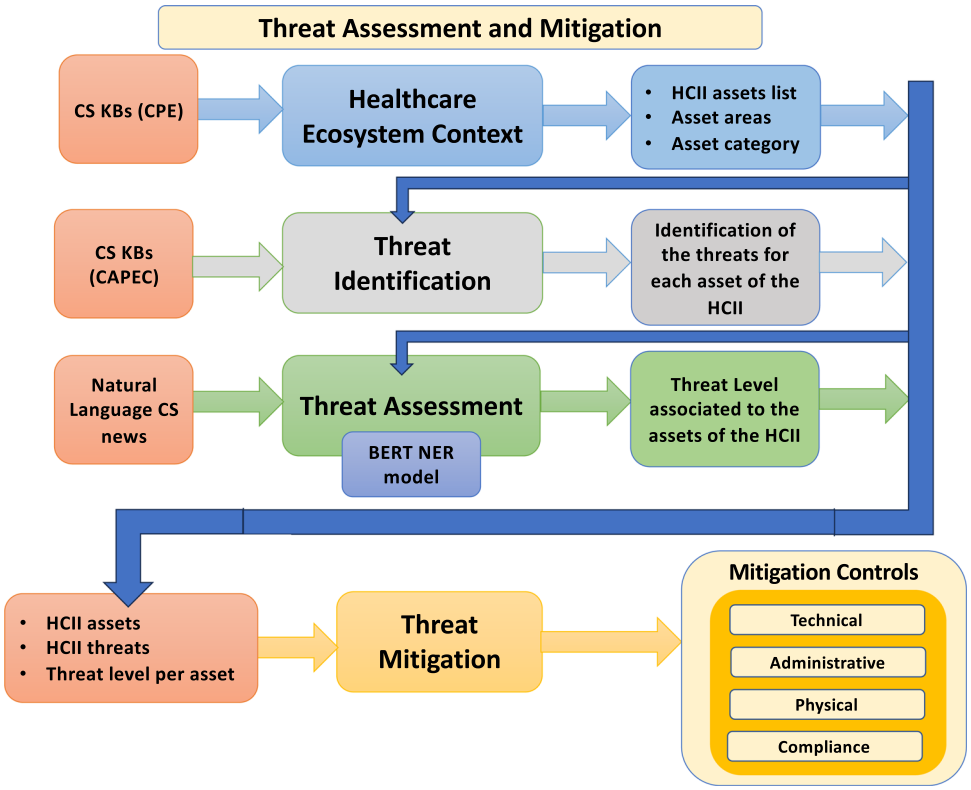

Fig. 2 Pipeline and physical view of the proposed threat assessment and mitigation methodology

4.2 CS news posts

A CS news posts collection has been used for creating both the annotated NER corpus, and to provide a NL corpus for the extraction of the information required for the threat level evaluation. This corpus has been extracted from The Hacker News website,[5] a CS news platform that attracts over 8 million readers monthly. This website is daily updated and contains tons of documents, describing attacks, threats, vulnerabilities, and other CS topics. We developed a web crawler and scraper for this website, which retrieves, extracts, collects and normalises only the text of each posted news. The scraping task is performed bi-weekly, making this dataset constantly updated and increasing its size. The dataset is currently (at the date of July 31st, 2023) counting 3,393,368 tokens and 133,637 sentences, extracted from 6774 news articles of the website.

In previous experiments described in [13], a training set and a test set for the NER module were created, randomly selecting a subset of the dataset and then using a semisupervised annotation methodology based on an iterative application of DS and AL, incrementally described and improved in [12, 13, 46, 47]. In summary, this annotation method requires a preliminary Distantly Supervised annotation of a small subset of the corpus (exploiting the domain KBs), which is then manually reviewed by a human. The obtained data is used to train a ML model (the same NLM used also as NER module, described in next Sect. 4.3), which can be used to automatically annotate another subset of the corpus. This new data annotated by the ML model is used to repeat the same procedure: is DS annotated, and manually reviewed by humans. Finally, the resulting annotated data is added to the data annotated in the previous iteration, and the obtained extended dataset is used to retrain the ML model with a higher precision, thanks to the larger dataset available. The new trained ML model is leveraged for the annotation of a new subset of data, and the same procedure is iteratively repeated, until the whole training set is annotated. This approach reduces the human effort required for the annotation, exploiting both domain KBs and an increasingly improved ML model, making the task of human annotators a simple review.

We applied this annotation method, adapting it to the CS domain by exploiting the KBs described in Sect. 4.1 for the distantly supervised annotation phase, as well as the NER model described in next Sect. 4.3 for the active learning phase. Using this annotation approach, we increased the size of these NER training set with respect to the previous experiments, thanks to the larger number of news collected from the web, with the purpose of improving the performances of the NER module. We used 1,002 news from the dataset to realise the annotated NER training and test sets.

[5] https://thehackernews.com.

Table 4 Hacker news datasets features

Dataset	News count	Word count
The Hacker news dataset	6,674	3,393,368
NER training set	918	453,021
NER test set	84	39,811
Threat level (TL) dataset	5672	2,900,536

The remaining CS news of the dataset (hereinafter called *Threat Level (TL)* dataset) are used to implement the corpus for the threat level assessment and mitigation, used to identify the sentences where the assets and corresponding threats are mentioned and to calculate the occurrence percentages for the assignment of the threat level. The details of these datasets are reported in the next Table 4.

The datasets will be shared publicly on the SoBigData research infrastructure,[6] in the SoBiogData Catalogue Section, where datasets and methods are made available for the scientific community.

4.3 Neural language models

The NLP NER module, which extracts the mentions of assets and threats from the CS natural language corpus, leverages a BERT-based NLM. We considered NLMs pretrained on large CS-domain corpora, with the purpose of improving the results of the downstream tasks applied to the same domain [49, 50]. To this end, in previous experiments described in [12, 13], we compared the performances of SecBERT,[7] CyNER,[8] and a general domain BERT model [10] (the first two NLMs are pretrained on closed-domain CS NL document corpora).

In detail, we first tested a CS closed-domain pretrained BERT model named SecBERT. This model is based on BERT-Base architecture, which is formed by 12 attention heads, 6 hidden layers and an hidden size equal to 768. The pretraining of SecBERT was performed on a very large CS document collection, which corpus includes: (i) APTnotes,[9] a collection of publicly-available papers and blogs (sorted by year) related to malicious campaigns, activity, or software that have been associated with vendor-defined APT (Advanced Persistent Threat) groups and/or tool-sets; (ii) the text extracted from the website included in Stucco-Data,[10] a repository that keeps a list of the data sources that are potentially relevant to cyber security and the source for the web site to make the data sources easy to read (including the texts from CPE, CVE and other databases, as well as blogs, forums, bulletin boards, etc.); (iii) a corpus of corpus of 1000 English news articles from 2017 to 2019 used for CASIE project [45]; (iv) the datasets of SemEval 2018 Task 8 SecureNLP [51], a shared task on semantic extraction from CS reports. SecBERT also includes a CS-specific masking strategy for the pretraining, which allows the model to better focus the attention on the CS-domain words.

The SecBERT model has been fine-tuned on the CS NER task, following the purposes of the proposed approach. In detail, using the NER training set described in Sect. 4.2, the model is trained to extract two classes of named entities: *assets* and *threats*.

As mentioned above, the other NLM tested as NER module is CyNER [43], which is based on an XLM RoBERTa-large NLM [52], specifically pretrained on unstructured thread reports and fine-tuned on CS NER task. Furthermore, it also integrates different additional modules for extracting CS entities, such as regular expressions and KBs, ML model for generic domain entities and a Flair-based [53] NER model. The selection of the entities is based on a priority merging technique. This model is already fine-tuned on the CS NER task and is able to recognise several named entity classes. Among them, the assets and threats necessary for the threat assessment respectively correspond to *indicator* and *malware* CyNER types in the first case, and to *system* and *organisation* CyNER entity type in the latter case.

We compared in our previous works [12, 13] the performances of the fine-tuned SecBERT and the CyNER models with a baseline fine-tuned BERT model, using the NER test set previously described in Sect. 4.2. Following the obtained results, we selected SecBERT as NER module for our threat assessment approach.

The NER model is used in conjunction with the KBs to identify the sentences of the TL Dataset where there is a mention of both an asset and a threat, allowing us to perform the proposed threat assessment methodology, obtaining the threat level based on the calculation of the percentage of occurrences of that threats.

4.4 Implementation: tools and resources

The implementation of the proposed approach requires CPE and CAPEC KBs described in Sect. 4.1 during the *Healthcare Ecosystem Context* and *Threat Identification* phases. The analysis and processing of the KBs can be supported by the web tools available on the respective websites. Moreover, the KBs can be downloaded in various formats and locally processed with custom scripts.

The *Threat Assessment* phase requires a large CS NL document collection, as the one described in Sect. 4.2. In our case, we developed a web crawler and scraper for the Hacker News web site implemented in Python language, using the

[6] http://www.sobigdata.eu.

[7] https://github.com/jackaduma/SecBERT.

[8] https://github.com/aiforsec/CyNER.

[9] https://github.com/aptnotes/data.

[10] http://stucco.github.io/data/.

BeautifulSoup library.[11] Moreover, some additional open source tools and libraries can be used for the prepossessing of the textual data. In particular, we used Spacy,[12] an NLP Python library that provides tools for tokenisation, sentence splitting and other NLP preprocessing tasks. Spacy was also used in the distantly supervised annotation of the NER dataset described in Sect. 4.1.

The fine-tuning of the SecBERT NLM has been performed using the Huggingface Python library,[13] which implements a set of API for training and fine-tuning Transformer-based NLMs. Finally, we exploited a specifically developed Python scripts to calculate the percentage of occurrence for the threat level assessment phase.

5 Evaluation

This Section presents the evaluation of the proposed approach and the obtained results. We have tested the threat assessment and mitigation approach in real settings, exploiting three different pilot studies provided by Fraunhofer Institute for Biomedical Engineering (IBMT),[14] a leading research institute in this medical engineering with clinical and industrial applications. The pilots described below a set of assets of healthcare ecosystem scenario.

5.1 Healthcare ecosystem scenario and assets identification

Fraunhofer IBMT has developed many solutions related to the healthcare ecosystem, which allowed us to provide a pilot environment to the NLP machine learning model. We assembled the following group of assets which describes different type of applications tested in pilots: *Implantable Medical Devices*, *Wearables*, and *Biobank*. As a result, the pilot provides a wide range of the assets, described in detail below and listed with the corresponding features in the next Tables 5, 6 and 7.

5.1.1 Implantable medical devices

We use the technological platform for programmable active implants developed during INTAKT project [54] to represent a use-case of a patient that suffers from tetraplegia. The implantable medical devices (IMD) are designed to help with restoring the grasping function of the upper limbs. There is a central control unit that manages the network of IMDs and allows external devices to communicate with the system, e.g. hospital IT infrastructure. The device is located on their body (e.g. belt). Firmware on the implants and the central control unit is implemented in C programming language using SimpleLink SDK [15] version 3.10.00.11. They run FreeRTOS [16] version 7.3.0 as a real-time operating system. In our setup, a doctor's computer, which connects with the central control units, runs on Windows 10 operating system and has a software installed to communicate with the system. The purpose of the software is to adjust different parameters on the system, such as stimulation patterns. The software is implemented in C++ programming language with usage of Qt6 framework. The doctor's PC is connected to two different servers running Ubuntu OS. One is a database server hosting a PostgreSQL database and the other one is Xploit Application Server, which is described in more detail in Sect. 5.1.2, where Wearables pilot is described.

Figure 3 depicts the IMD pilot. It consists of the following assets:

- Implant
- Central Control Unit
- Doctor's PC
- Unix Server with a Database

5.1.2 Wearables

The part of the pilot related to wearables focuses on a platform for patient monitoring and symptom reporting (s. Fig. 4). The platform was developed by Fraunhofer IBMT in the context of the COVID-19 pandemic. It addresses relevant cybersecurity issues in this typical scenario for e-health solutions with wearables. The platform consists of the commercial smartwatch ScanWatch by Withings in combination with the app Corona Diary by Fraunhofer IBMT, which was used in a clinical pilot project by Fraunhofer IBMT and the University Hospital of Saarland to collect self-reported symptom data from COVID-19 patients in home quarantine for research purposes. The pilot for wearables consists of the following assets:

- XplOit Application
- XplOit Application Server
- Withings Server
- Corona Diary App
- ScanWatch
- XplOit Triple Store Virtuoso
- XplOit Triple Store
- XplOit Document-oriented database

[11] https://www.crummy.com/software/BeautifulSoup/bs4/doc/.

[12] https://spacy.io.

[13] https://huggingface.co.

[14] https://www.ibmt.fraunhofer.de/en/ibmt-institute-profile/ibmt-history.html.

[15] https://www.ti.com/tool/download/SIMPLELINK-CC13X0-SDK/3.10.00.11.

[16] https://www.freertos.org/.

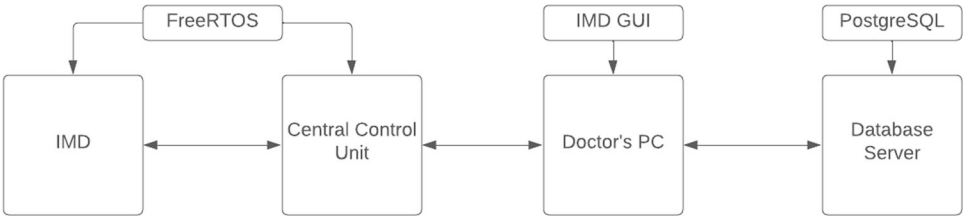

Fig. 3 Overview of the implantable medical pilot

Table 5 Details of the assets of the implantable medical devices pilot (use case 1)

Asset name	Vendor	Product name	Category	Area
Implantable device	Texas Instrument	SimpleLink	Embedded device	1
FreeRTOS	Amazon	FreeRTOS	Operating System	3
Central Control Unit	Texas Instrument	SimpleLink	Embedded device	2
FreeRTOS	Amazon	FreeRTOS	Operating System	3
Doctor's PC	Microsoft	Windows 10	Operating System	3
Database server	Ubuntu	Ubuntu Linux	Operating System	4
Database	PostrgreSQL GDG	PostgreSQL	Database	3
IMD GUI	Fraunhofer IBMT	IMD GUI	Software	3

– Operating System Ubuntu

The next Fig. 4 shows an overview of the wearables pilot, while Table 6 summarises the details of these assets.

5.1.3 Biobank

Fraunhofer IBMT collects and manages important biorepositories and provides human biomaterial for research purposes. It also collects and stores human samples from specific donor cohorts to exposure to contaminants in the environment on behalf of the German Environment Agency (UBA). For the environmental study "Environmental Survey for Children and Adolescents" (2014–2017) conducted by UBA, Fraunhofer IBMT developed the sample management system UBA-PVS to collect, process, store, and manage samples and associated data from around 2500 participants and more than 70,000 samples. UBA-PVS represents the information system for the pilot on cybersecurity in biobanks. The pilot for biobanks consists of the following assets:

– UBA-PVS Application
– Operating System Ubuntu
– UBA-PVS Database
– UBA-PVS Web Application Framework: Hibernate, PrimeFaces, Spring
– UBA-PVS Application Server
– Operating System Ubuntu

The next Table 7 summarises the assets described above with the details about their name, vendor, product name and type, also including the corresponding areas as defined in previous Sect. 3.1.2.

5.2 Experiments, results and discussion

The experimental assessment tested the capability of the proposed approach on the Fraunhofer IBMT healthcare ecosystems described in previous Sect. 5.1, included in the pilots of the EU-funded AI4HEALTHSEC project. For this purpose, we firstly applied the fine-tuned NER model to the TL dataset, after previously preprocessing the text with a sentence splitting software, in order to extract the mentions of threats and assets that are present in the same sentence.

The number of sentences, the assets (from the IBMT pilots) and threats in the same sentences found by the NER module are reported in Table 8. Some examples of sentences found by the proposed approach, which contain a mention of both assets (from the considered pilots) and corresponding threats are reported below (the asset entities are in blue, and the threat entities are in red):

– *In one unsuccessful incident targeting an unspecified customer, the actor targeted the Zoho ManageEngine ADSelfService Plus service running on an Apache Tomcat server to trigger the execution of suspicious commands pertaining to process enumeration and network connectivity, among others.*
– *This flaw allows a local attacker with a user account on the system to gain access to out-of-bounds memory, leading to a system crash or a privilege escalation threat, Red Hat said in an advisory published on February 22, 2022, and similar alerts have been released by Debian, Oracle Linux, SUSE, and Ubuntu.*
– *CVE-2022-22963 - VMware Tanzu Spring Cloud Function Remote Code Execution Vulnerability.*

Table 6 Details of the assets of the wearable pilot (use case 2)

Asset name	Vendor	Product name	Category	Area
Xploit Application Server	Apache Software Foundation	Apache Tomcat Webserver	Application Server	3
XplOit document-oriented DB	MongoDB Inc	MongoDB CE	Database	3
Xploit Triple Store Ontology DB	OpenLink Software	Openlink Virtuoso	Database	3
ScanWatch	Withings	Withings Scan Watch	Smart Watch OS	3
Corona Diary App	Fraunhofer IBMT	Corona Diary App	Android App	3
Withings Server API	Withings	Withings API	API	3
XplOit Application	Fraunhofer IBMT	XplOit	Web Application	3
Server OS	Ubuntu	Ubuntu Linux	OS	4

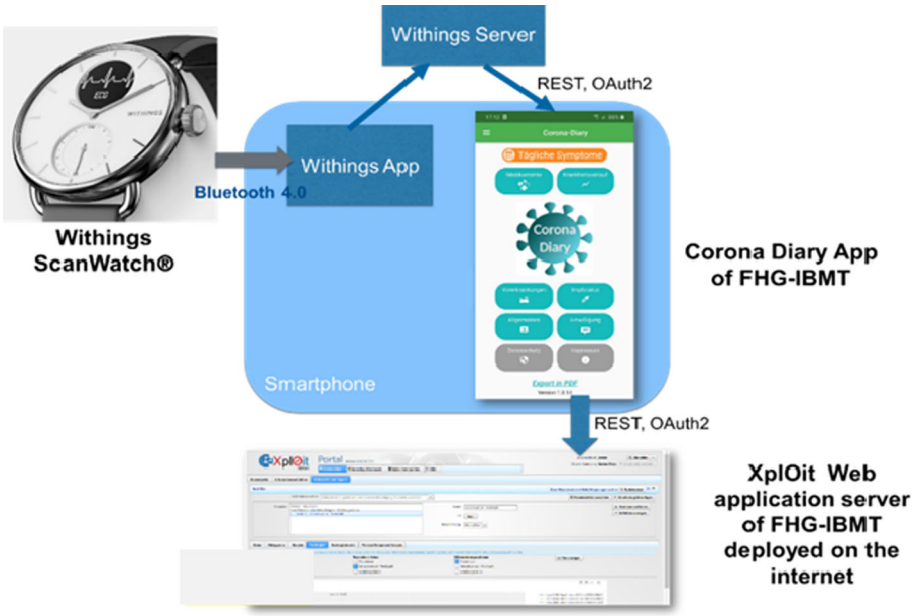

Fig. 4 Overview of the wearables pilot

Table 7 Details of the assets of the biobank pilot (use case 3)

Asset name	Vendor	Product name	Category	Area
UBA-PVS Application	Fraunhofer IBMT	UBA-PVS	Web Application	3
UBA-PVS Application Server	Apache Software Foundation	Apache Tomcat Webserver	Application Server	3
UBA-PVS Database	PostrgreSQL GDG	PostgreSQL	Database	3
UBA-PVS Web Application Framework 1	VMWare	Spring Framework	Web Application Framework	3
UBA-PVS Web Application Framework 2	PrimeTek Informatics	Primefaces	Web Application Framework	3
UBA-PVS Web Application Framework 3	JBoss (Red Hat)	Hibernate	Web Application Framework	3
Server OS	Ubuntu	Ubuntu Linux	OS	4

– *In February 2022, HP detailed a social engineering attack using fake Windows 11 upgrade installers to trick Windows 10 users into downloading and executing the textcolorredmalware.*

After extracting the entities and the sentences containing at least both an asset and a threat, it is possible to create the threat occurrence table for each pair asset/threat mentioned in the same sentence. In this case, we implemented a Python script that is able to calculate the percentages of occurrences

Table 8 Sentences with threats and assets entities from the IBMT pilots extracted from the TL dataset

Assets count	Threats count	Sentence count
1585	1334	1288

of each pair with respect to the total number of sentences where the threats are mentioned for each asset. Therefore, following the ranges of the percentage of occurrence previously described and shown in Table 3, it is possible to assign the level of every threat of each asset of the pilots. The threat levels for the assets of each uses case are respectively reported in Table 9 (the assets of the pilots not included in the Table have not been found in the current TL dataset).

The obtained results highlight the assets and the corresponding threat levels for each pilot environment tested, making it possible to understand where it is more necessary to focus the required mitigation actions. In particular, we can see that, in the case of the Implantable Medical Device (use case 1), the higher level threats are related to Windows 10, where several threats have high and medium levels. The PostgreSQL asset also shows some threats with medium and high levels. On the other hand, although several threats have been found for Ubuntu Linux (used in all three use cases), their level ranges from low to very low. The Wearable pilot (use case 2) has the potentially more dangerous threats related to the PostgreSQL and Apache Tomcat assets, but, in general, the threat levels of the assets involved and the number of pairs asset/threats found in this use case are lower with respect to the use case 1. Finally, only one high-level threat has been found in the Biobank pilot (use case 3), which is related to the execution of suspicious or arbitrary code exploiting the Apache Tomcat asset. In general, the assets of use case 3 show a limited number of threats, with low or very low levels in most cases.

As explained, the assignment of the threat level allows to select the most appropriate actions to improve the CS level of the considered HCII, as well as to select the most appropriate mitigation actions. We also remark that, as new and updated textual data are obtained (the scraping of the Hacker News website is performed monthly), the proposed approach is reapplied to a larger and updated dataset and the percentages of assets/threats occurrences can be consequently updated, as well as new assets/threats pairs could be found and added to the threat occurrence table.

The obtained results demonstrated that the proposed approach can exploit NL CS documents to calculate the level associated with the threats extracted from CS narrative documents. This information can be mapped to the assets in the Heath Care Information Infrastructures and in their supply chains. In this way, it was possible to identify a significant number of threats for a set of assets involved in the HCII and evaluate their corresponding level.

Moreover, it is worth noting that the values of the metrics of the obtained results suggest that this method is exploitable to develop the overall CS situational awareness, supporting the monitoring and the prevention of CS incidents in the HCIIs in real-world environments. In particular, the obtained results provide information that helps to evaluate and select which assets and threats require the most crucial mitigation actions. Threat mitigation attempts to look at various controls including functionalities relating to corrective, detective and preventive measures.

In the cases of the considered pilots, arbitrary code injection is ranked as high threat level for a number of assets of Fraunhofer IBMT relating to Implantable Medical Devices, Wearables, and Biobank. Such a threat impacts on different vendor products, i.e. Apache Tomcat, Windows 10, and Spring framework, which are managing the Fraunhofer IBMT medical devices. This threat can allow attackers to take control of the medical devices. It is necessary to consider preventative measures to mitigate this threat. Apache Tomcat requires to install a source code patch specific to the version to mitigate this threat which is available from the website. Specifically, if the FROM authentication is used by the root web application, then the patch needs to be installed. Spring Framework specific versions such as 5.3.0 to older ones are affected by this threat which allows to execute remote code on the applications running through this framework specifically with the mapping functionalities. Similar to Tomcat, it is also necessary to update the patch from the Spring framework and Windows system, to mitigate this threat. Elevate privileges in Windows is also ranked as medium threat level, which allows threat actors to elevate process privileges on medical devices running through Windows systems. However, this threat cannot target the latest Windows version and there is a need to install various security features such as Supervisor Mode Execution Prevention (SMEP) and Virtualisation-Based Security (VBS) to mitigate this threat.

6 Conclusion, limitations and future work

The complexity of the healthcare sector is now significantly increased due to the rapid digitisation and adoption of more connected medical devices. Cyber attacks within the sector are also increasing, therefore it is necessary to ensure secure healthcare service delivery. This paper contributes towards this direction by proposing a novel threat assessment and management approach. Threat identification and management is one of the critical steps towards security and key threat-related information is available through the unstructured natural language documents.

The proposed method assesses a specific threat exploiting assets of healthcare infrastructure. In particular, after a preliminary identification of the healthcare ecosystem context, where the services and the assets are identified and categorised, the method includes a threat identification and assessment phase, based on the percentage of occurrence in NL documents containing CS topics and extracted from online sources, such as news sites, social media, forums, and others. The mentions of the threats and assets are extracted

Table 9 Asset/threat pairs, corresponding threat level, and use cases where the assets are included

Assets	Threat	Threat level	Use cases
Apache Tomcat	Execution of suspicious/arbitrary code or code injection	High	2, 3
Apache Tomcat	Elevate privileges	Very Low	2, 3
Apache Tomcat	Bypass authentication	Very Low	2, 3
Apache Tomcat	Run unsigned/unauthorised software	Very Low	2, 3
Apache Tomcat	Package highjacking	Very Low	2, 3
Apache Tomcat	Malware propagation	Very Low	2, 3
Apache Tomcat	Remote access/backdoor	Very Low	2, 3
MongoDB	Execution of suspicious/arbitrary code or code injection	Medium	3
MongoDB	Denial of Service	Very Low	3
MongoDB	Package highjacking	Low	3
Ubuntu Linux	Execution of suspicious/arbitrary code or code injection	Low	1, 2, 3
Ubuntu Linux	Elevate privileges	Low	1, 2, 3
Ubuntu Linux	Run unsigned/unauthorised software	Very Low	1, 2, 3
Ubuntu Linux	Denial of Service	Low	1, 2, 3
Ubuntu Linux	Package highjacking	Very Low	1, 2, 3
Ubuntu Linux	Secure boot bypass	Very Low	1, 2, 3
Ubuntu Linux	Information disclosure/data theft	Very Low	1, 2, 3
Ubuntu Linux	Malware propagation	Very Low	1, 2, 3
Ubuntu Linux	Remote access/backdoor	Very Low	1, 2, 3
Windows 10	Execution of suspicious/arbitrary code or code injection	High	1
Windows 10	Elevate privileges	Medium	1
Windows 10	Bypass authentication	Very Low	1
Windows 10	Run unsigned/unauthorised software	Very Low	1
Windows 10	Denial of service	High	1
Windows 10	Package highjacking	Very Low	1
Windows 10	Secure boot bypass	Very Low	1
Windows 10	Information disclosure/data theft	Medium	1
Windows 10	Malware propagation	Medium	1
Windows 10	Remote access/backdoor	High	1
PostgreSQL	Execution of suspicious/arbitrary code or code injection	Medium	1, 2
PostgreSQL	Elevate privileges	Very Low	1, 2
PostgreSQL	Bypass authentication	Very Low	1, 2
PostgreSQL	Denial of service	Very Low	1, 2
PostgreSQL	Package highjacking	Very Low	1, 2
PostgreSQL	Information disclosure/data theft	High	1, 2
PostgreSQL	Malware propagation	Low	1, 2
PostgreSQL	Remote access/backdoor	Medium	1, 2
Spring	Execution of suspicious/arbitrary code or code injection	High	2
Spring	Run unsigned/unauthorised software	Very Low	2
Spring	Denial of service	Very Low	2
Spring	Information disclosure/data theft	Very Low	2
Spring	Malware propagation	Low	2
Spring	Remote access/backdoor	Very Low	2
Primefaces	Remote access/backdoor	Very Low	2

by applying a Transformer-based NER module specifically fine-tuned for this task. Finally, a threat mitigation phase aims to mitigate the identified threats based on the informed decision making, considering the identified threats and their corresponding levels.

We applied the proposed approach to three real-world healthcare ecosystem scenarios provided by Fraunhofer IBMT, namely Implantable Medical Devices, Wearables and Biobank pilots. These use cases are included in the pilot studies of the AI4HEALTHSEC EC-funded project, demonstrating its effectiveness in identifying the pairs assets/threats and calculating the threat level and the corresponding potentially involved assets of the HCII, suggesting in this way the more appropriate mitigation actions.

Although the proposed approach can work with any kind of natural language text (not requiring specific text formats), a limitation is related to the acquisition of large NL corpora, where reports of CS threats are described. Larger datasets can provide a wider coverage of assets and threats and more accurate information. This kind of data is largely available on the Internet, but in many cases, specific web crawlers and web scrapers must be developed, as well as it could be necessary to rely on tools provided by the owner of the website. An example of this issue is Twitter, which allows users to acquire the tweets data using their closed API, with some limitations, such as the impossibility of retrieving tweets more than seven days old.

Another limit is related to the language of the NL corpus. The NLP NER model is currently trained only on English and, in the case of the need of processing CS NL documents in other languages, it is necessary to leverage an NLM pretrained for that specific language and fine-tuned on annotated NER task. Even if cross-language models and multilingual approaches reported recently good performances in literature [55, 56], a training set must also be annotated in the new language, requiring a further effort. As explained in Sect. 4.3, a domain-specific NLM can improve the performances of the NER module and currently, in our knowledge, there are no domain-specific NLM for CS domain pretrained in languages different from English.

Concerning the applicability of our model to other businesses and contexts, another limit is the need of domain-specific KBs, to correctly model the assets of the considered use case, during the preliminary phases of the proposed approach. This issue could limit its applicability in some domains, where many custom assets are involved. Moreover, to improve the performances of the NER module, a domain-pretrained model should be used.

The model is also continuously updated, by a monthly crawling of CS news from the Internet, providing constantly updated information that can be used to improve the obtained results. Moreover, in the future, different types of NL CS datasets can be included to further enlarge the information source adopted to evaluate the threat level and the information about the threats affecting the assets of the HCII, such as CS topic social media posts, CS forums posts and others NL data publicity available. In this way, the threat level identification will be more trustable, relying on larger datasets. Other improvements can be obtained by adopting more complex and recent NLMs, such as ChatGPT 3 or later versions, as w2ell as to implement and integrate into our pipeline a ML-based Relation Extraction method, capable of automatically detecting and classifying the relations between the threat and the assets.

Finally, to the end of further mitigating the possible contribution of false positives and negatives found by the ML NLP module to the threat level, in addition to a constant increasing of the NL corpora size, we could integrate the threat prioritisation obtained by our method also with the information available in the CS KBs, correcting the threat level with a weight extracted from the KBs.

Acknowledgements This work is supported by European Union—NextGenerationEU—National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR)—Project: "SoBigData.it—Strengthening the Italian RI for Social Mining and Big Data Analytics"—Prot. IR0000013—Avviso n. 3264 del 28/12/2021, and by the European Commission, grant number 883273, AI4HEALTHSEC—A Dynamic and Self-Organised Artificial Swarm Intelligence Solution for Security and Privacy Threats in Healthcare ICT Infrastructures, and by European Union's Horizon Europe research and innovation program under grant agreement No 101120779, CyberSecDome—An innovative Virtual Reality-based intrusion detection, incident investigation and response approach for enhancing the resilience, security, privacy and accountability of complex and heterogeneous digital systems and infrastructures.

Funding Open access funding provided by Consiglio Nazionale Delle Ricerche (CNR) within the CRUI-CARE Agreement.

Data availability The datasets generated during and/or analysed during the current study are available in the *SoBigData* repository, http://www.sobigdata.eu/.

Declarations

Conflict of interest The authors participated in project funded by the European Commission AI4HEALTHSEC—A Dynamic and Self-Organised Artificial Swarm Intelligence Solution for Security and Privacy Threats in Healthcare ICT Infrastructures (grant number 883273). The article describes some of the results of this project.

Ethical standards This article does not contain any studies with human participants and/or animals performed by any of the authors.

References

1. Rees, D.: Cyber attacks in healthcare: the position across Europe (2021). https://www.pinsentmasons.com/out-law/analysis/cyber-attacks-healthcare-europe
2. McKee, D., Laulheret, P.: McAfee Enterprise ATR uncovers vulnerabilities in globally used B. Braun infusion pump (2021). https://www.mcafee.com/blogs/enterprise/mcafee-enterprise-atr/mcafee-enterprise-atr-uncovers-vulnerabilities-in-globally-used-b-braun-infusion-pump/
3. Institute, P.: Sixth annual benchmark study on privacy & security of healthcare data. Tech. rep, Ponemon Institute (2016)
4. Islam, S., Papastergiou, S., Mouratidis, H.: A dynamic cyber security situational awareness framework for healthcare ICT infrastructures. In: PCI 2021: 25th Pan-Hellenic Conference on Informatics, pp. 334–339. ACM, Volos, Greece (2021). https://doi.org/10.1145/3503823.3503885
5. Tikhomirov, M., Loukachevitch, N.V., Sirotina, A., Dobrov, B.V.: Using BERT and augmentation in named entity recognition for cybersecurity domain. In: Natural Language Processing and Information Systems—25th International Conference on Applications of Natural Language to Information Systems, NLDB 2020, vol. 12089, pp. 16–24. Springer, Saarbrücken, Germany (2020). https://doi.org/10.1007/978-3-030-51310-8_2
6. Ma, P., Jiang, B., Lu, Z., Li, N., Jiang, Z.: Cybersecurity named entity recognition using bidirectional long short-term memory with conditional random fields. Tsinghua Sci. Technol. **26**(3), 259 (2021). https://doi.org/10.26599/TST.2019.9010033
7. Zhou, S., Liu, J., Zhong, X., Zhao, W.: Named entity recognition using BERT with whole world masking in cybersecurity domain. In: 2021 IEEE 6th International Conference on Big Data Analytics (ICBDA), pp. 316–320. IEEE, Xiamen, China (2021). https://doi.org/10.1109/ICBDA51983.2021.9403180
8. Chen, Y., Ding, J., Li, D., Chen, Z.: Joint BERT model based cybersecurity named entity recognition. In: 2021 The 4th International Conference on Software Engineering and Information Management. Association for Computing Machinery, Yokohama, Japan, 2021, pp. 236-242. ICSIM (2021). https://doi.org/10.1145/3451471.3451508
9. Gao, C., Zhang, X., Liu, H.: Data and knowledge-driven named entity recognition for cyber security. Cybersecurity **4**(1), 1 (2021). https://doi.org/10.1186/s42400-021-00072-y
10. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: BERT: Pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vol. 1, pp. 4171–4186. ACL, Minneapolis, Minnesota (2019). https://doi.org/10.18653/v1/N19-1423
11. Islam, S., Papastergiou, S., Silvestri, S.: Cyber threat analysis using natural language processing for a secure healthcare system. In: 2022 IEEE Symposium on Computers and Communications (ISCC), pp. 1–7. IEEE, Rhodes, Greece (2022). https://doi.org/10.1109/ISCC55528.2022.9912768
12. Islam, S., Papastergiou, S., Silvestri, S.: Cyber threat analysis using natural language processing for a secure healthcare system. In: 2022 IEEE Symposium on Computers and Communications (ISCC), pp. 1–7. IEEE, Rhodes, Greece (2022). https://doi.org/10.1109/ISCC55528.2022.9912768
13. Silvestri, S., Islam, S., Papastergiou, S., Tzagkarakis, C., Ciampi, M.: A machine learning approach for the NLP-based analysis of cyber threats and vulnerabilities of the healthcare ecosystem. Sensors **23**(2), 651 (2023). https://doi.org/10.3390/s23020651
14. Goud, N.: Malware and ransomware attack on medical devices (2017). https://www.cybersecurity-insiders.com/malware-and-ransomware-attack-on-medical-devices/
15. Argaw, S.T., Troncoso-Pastoriza, J.R., Lacey, D., Florin, M., Calcavecchia, F., Anderson, D., Burleson, W.P., Vogel, J., O'Leary, C., Eshaya-Chauvin, B., Flahault, A.: Cybersecurity of hospitals: discussing the challenges and working towards mitigating the risks. BMC Med. Inform. Decis. Mak. **20**(1), 146 (2020). https://doi.org/10.1186/s12911-020-01161-7
16. Rios, B., Butts, J.: Security evaluation of the implantable cardiac device ecosystem architecture and implementation interdependencies (2017). https://a51.nl/sites/default/files/pdf/Pacemaker%20Ecosystem%20Evaluation.pdf
17. Nifakos, S., Chandramouli, K., Nikolaou, C.K., Papachristou, P., Koch, S., Panaousis, E., Bonacina, S.: Influence of human factors on cyber security within healthcare organisations: a systematic review. Sensors **21**(15), 5119 (2021). https://doi.org/10.3390/s21155119
18. Snell, E.: Hacking still leading cause of 2015 health data breaches. Health IT Security (2015)
19. Halperin, D., Heydt-Benjamin, T.S., Ransford, B., Clark, S.S., Defend, B., Morgan, W., Fu, K., Kohno, T., Maisel, W.H.: Pacemakers and implantable cardiac defibrillators: Software radio attacks and zero-power defenses. In: 2008 IEEE Symposium on Security and Privacy (sp 2008), pp. 129–142. IEEE (2008)
20. Storm, D.: Medjack, hackers hijacking medical devices to create backdoors in hospital networks. Comput. World **8**, 42 (2015)
21. CIS. Cyber attacks: In the healthcare sector. [online]. https://www.cisecurity.org/blog/cyber-attacks-in-the-healthcare-sector. Accessed 25 May 2023
22. Islam, S., Papastergiou, S., Kalogeraki, E.M., Kioskli, K.: Cyber-attack path generation and prioritisation for securing healthcare systems. Appl. Sci. **12**(9), 4443 (2022)
23. Shevchenko, N.: Threat modeling: 12 available methods (2018). https://insights.sei.cmu.edu/blog/threat-modeling-12-available-methods/
24. Microsoft. Stride model (2022). https://learn.microsoft.com/en-us/azure/security/develop/threat-modeling-tool-threats#stride-model. Accessed 22 Sept 2023
25. Schneier, B.: Modeling security threats (1999). https://www.schneier.com/academic/archives/1999/12/attack_trees.html
26. Alwaheidi, S., Islam, M.K.S.: Data-driven threat analysis for ensuring security in cloud enabled systems. Sensors **22**(15), 5726 (2022). https://doi.org/10.3390/s22155726
27. Owasp. Owasp threat dragon. https://owasp.org/www-project-threat-dragon
28. Rak, M., Salzillo, G., Granata, D.: Esseca: an automated expert system for threat modelling and penetration testing for IoT ecosystems. Comput. Electr. Eng. **99**, 107721 (2022). https://doi.org/10.1016/j.compeleceng.2022.107721
29. Threatmodeler. Threat modeling for healthcare organizations (2023). https://threatmodeler.com/threat-modeling-for-healthcare-organizations
30. Omotosho, A., Haruna, B.A., Olaniyi, O.M.: Threat modeling of internet of things health devices. J. Appl. Secur. Res. **14**(1), 106 (2019). https://doi.org/10.1080/19361610.2019.1545278
31. Almohri, H., Cheng, L., Yao, D., Alemzadeh, H.: On threat modeling and mitigation of medical cyber-physical systems. In: 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE), pp. 114–119 (2017). https://doi.org/10.1109/CHASE.2017.69

32. Bharathi, V., Kumar, C.V.: A real time health care cyber attack detection using ensemble classifier. Comput. Electr. Eng. **101**, 108043 (2022)
33. Yeboah-Ofori, A., Mouratidis, H., Ismai, U., Islam, S., Papastergiou, S.: Cyber supply chain threat analysis and prediction using machine learning and ontology. In: Artificial Intelligence Applications and Innovations—17th IFIP WG 12.5 International Conference, AIAI 2021, vol. 627, pp. 518–530. Springer, Hersonissos, Crete, Greece (2021). https://doi.org/10.1007/978-3-030-79150-6_41
34. Haque, N.I., Rahman, M.A., Shahriar, M.H., Khalil, A.A., Uluagac, A.S.: A novel framework for threat analysis of machine learning-based smart healthcare systems, CoRR **abs/2103.03472** (2021)
35. Radanliev, P., Roure, D.D.: Advancing the cybersecurity of the healthcare system with self-optimising and self-adaptative artificial intelligence (part 2). Heal. Technol. **12**, 923–929 (2022)
36. Zong, S., Ritter, A., Mueller, G., Wright, E.: Analyzing the perceived severity of cybersecurity threats reported on social media. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vol. 1, pp. 1380–1390. Association for Computational Linguistics, Minneapolis, Minnesota (2019). https://doi.org/10.18653/v1/N19-1140
37. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I.: Attention is all you need. In: Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, pp. 5998–6008. Long Beach, CA, USA (2017)
38. Bayer, M., Kuehn, P., Shanehsaz, R., Reuter, C.: Cysecbert: A domain-adapted language model for the cybersecurity domain, CoRR **abs/2212.02974**. https://doi.org/10.48550/arXiv.2212.02974 (2022)
39. Ranade, P., Piplai, A., Joshi, A., Finin, T.: Cybert: Contextualized embeddings for the cybersecurity domain. In: 2021 IEEE International Conference on Big Data (Big Data), pp. 3334–3342 (2021). https://doi.org/10.1109/BigData52589.2021.9671824
40. Ameri, K., Hempel, M., Sharif, H., Lopez, J., Jr., Perumalla, K.: Cybert: cybersecurity claim classification by fine-tuning the bert language model. J. Cybersecur. Privacy **1**(4), 615 (2021). https://doi.org/10.3390/jcp1040031
41. Ameri, K., Hempel, M., Sharif, H., Lopez, J., Perumalla, K.: Design of a novel information system for semi-automated management of cybersecurity in industrial control systems. ACM Trans. Manag. Inf. Syst. **14**(1), 58 (2023). https://doi.org/10.1145/3546580
42. Aghaei, E., Niu, X., Shadid, W., Al-Shaer, E.: Securebert: a domain-specific language model for cybersecurity. In: Li, F., Liang, K., Lin, Z., Katsikas, S.K. (eds.) Secur. Privacy Commun. Netw., pp. 39–56. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-15255-9_5
43. Alam, M.T., Bhusal, D., Park, Y., Rastogi, N.: CyNER: a python library for cybersecurity named entity recognition, CoRR **abs/2204.05754**. https://doi.org/10.48550/arXiv.2204.05754 (2022)
44. Fujii, S., Kawaguchi, N., Shigemoto, T., Yamauchi, T.: Cyner: information extraction from unstructured text of CTI sources with noncontextual iocs. In: Cheng, C.M., Akiyama, M. (eds.) Adv. Inf. Comput. Secur., pp. 85–104. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-15255-9_5
45. Satyapanich, T., Ferraro, F., Finin, T.: CASIE: extracting cybersecurity event information from text. In: The Thirty-Fourth AAAI Conference on Artificial Intelligence, AAAI 2020, The Thirty-Second Innovative Applications of Artificial Intelligence Conference, IAAI 2020, pp. 8749–8757. AAAI Press, New York, NY, USA (2020)
46. Silvestri, S., Gargiulo, F., Ciampi, M.: Iterative annotation of biomedical NER corpora with deep neural networks and knowledge bases. Appl. Sci. **12**(12), 5775 (2022). https://doi.org/10.3390/app12125775
47. Aracri, G., Folino, A., Silvestri, S.: Integrated use of KOS and deep learning for data set annotation in tourism domain. J. Doc. (2023). https://doi.org/10.1108/JD-02-2023-0019
48. Ciampi, M., Sicuranza, M., Silvestri, S.: A privacy-preserving and standard-based architecture for secondary use of clinical data. Information **13**(2), 87 (2022). https://doi.org/10.3390/info13020087
49. Silvestri, S., Gargiulo, F., Ciampi, M.: Improving biomedical information extraction with word embeddings trained on closed-domain corpora. In: 2019 IEEE Symposium on Computers and Communications (ISCC)), pp. 1129–1134. IEEE (2019). https://doi.org/10.1109/ISCC47284.2019.8969769
50. Yu, X., Hu, W., Lu, S., Sun, X., Yuan, Z.: BioBERT based named entity recognition in electronic medical record. In: 2019 10th International Conference on Information Technology in Medicine and Education (ITME), pp. 49–52 (2019). https://doi.org/10.1109/ITME.2019.00022
51. Phandi, P., Silva, A., Lu, W.: SemEval-2018 task 8: Semantic extraction from CybersecUrity REports using natural language processing (SecureNLP). In: Proceedings of The 12th International Workshop on Semantic Evaluation, pp. 697–706. Association for Computational Linguistics, New Orleans, Louisiana (2018). https://doi.org/10.18653/v1/S18-1113
52. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V.: Roberta: A robustly optimized BERT pretraining approach, CoRR **abs/1907.11692**. http://arxiv.org/abs/1907.11692 (2019)
53. Akbik, A., Bergmann, T., Blythe, D., Rasul, K., Schweter, S., Vollgraf, R.: FLAIR: An easy-to-use framework for state-of-the-art NLP. In: Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (Demonstrations), pp. 54–59. Association for Computational Linguistics, Minneapolis, Minnesota, USA (2019). https://doi.org/10.18653/v1/N19-4010
54. Fraunhofer Institute for Integrated Circuits IIS: Project INTAKT. https://www.iis.fraunhofer.de/en/ff/sse/sensorsolutions/forschung/intakt.html. Accessed 18 Oct 2023
55. Guarasci, R., Silvestri, S., De Pietro, G., Fujita, H., Esposito, M.: Bert syntactic transfer: a computational experiment on Italian, French and English languages. Comput. Speech Lang. **71**, 101261 (2022). https://doi.org/10.1016/j.csl.2021.101261
56. Silvestri, S., Gargiulo, F., Ciampi, M., De Pietro, G.: Exploit multilingual language model at scale for ICD-10 clinical text classification. In: 2020 IEEE Symposium on Computers and Communications (ISCC), pp. 1–7 (2020). https://doi.org/10.1109/ISCC50000.2020.9219640