UNIVERSITÀ
DEGLI STUDI
FIRENZE

Università degli Studi di Firenze
Dipartimento di Ingegneria dell'Informazione (DINFO)
Corso di Dottorato in Information Engineering
Curriculum: Informatica

Settore Scientifico Disciplinare ING-INF/06

Designing and developing a dedicated Natural Language Processing Framework for Healthcare Information Technology Management and Assessment

Candidate
Dott. Ing. Alessio Luschi

Supervisors
Prof. Paolo Nesi

Prof. Ernesto Iadanza

PhD Coordinator
Prof. Fabio Schoen

---

ciclo XXXVI, 2020-2023

Università degli Studi di Firenze, Dipartimento di Ingegneria dell'Informazione (DINFO).

A me

Acknowledgments

I would like to acknowledge the efforts and input of my supervisors, Prof. Paolo Nesi and Prof. Ernesto Iadanza. Many thanks go to Prof. Farah Magrabi and her team at Macquaire University in Sidney (Australia) for providing the initial labelled dataset. I would like to thank also my closest friends, Giulio, Antongiulio, Federico, and Cristian. Finally, I want to express my deepest gratitude to my beloved Alice, whose unwavering support has been my guiding light throughout these challenging years. Her radiant smile and infectious joy have been my constant source of inspiration and strength. None of this would have been achievable without all of you by my side.

Abstract

The escalating complexity of the hospital environment, propelled by technological advancements, necessitates a comprehensive exploration of the integration and management of diverse tools and technologies in healthcare settings. In this context, digital solutions, including the Internet of Things, robotics, mobile apps, sensors, and Artificial Intelligence (AI), play pivotal roles in enhancing treatment efficacy, ensuring patient safety, and optimising resource utilisation. The proliferation of health technologies demands a robust strategy for investigating factors affecting patient safety, necessitating interventions grounded in evidence-based approaches. This thesis delves into the critical analysis of manufacturer-recommended maintenance practices, urging Clinical Engineers and Health Technology Management professionals to adopt evidence-based methods. Real-World Data emerges as a valuable resource, offering observational insights into the effectiveness and safety of health technologies, with implications for regulatory decision-making, compliance with the EU Medical Device Regulation, and post-market surveillance. Addressing these challenges, this manuscript promotes the application of semantic ontologies to standardise data and enhance communication across healthcare systems. It highlights the role of semantic ontologies in managing the complexity of healthcare facilities, facilitating communication among various roles, and bridging gaps in data standardisation. The central focus of the work is developing a framework employing Natural Language Processing, Deep Neural Networks, and Explainable AI to extract and classify adverse events related to Health Information Technologies. Leveraging records from the US Manufacturer and User Device Experience database, the framework aims to provide a novel approach for obtaining Real-World Evidence in Clinical Engineering fields, including Evidence-Based Maintenance, Health Technology Management, and Assessment.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

The hospital environment is becoming more and more complex as technological development is advancing [88]. Nowadays, healthcare facilities incorporate different tools and technologies for empowering the efficacy and efficiency of health treatments, and for minimizing the obstacles about accessibility and cooperation, as well as strengthening patient safety, productivity, and quality of the working environment, while preserving cost-effectiveness. Digital solutions which support services and resources are being introduced in this scenario: Internet of Things (IoT), robotics, mobile apps, sensors, and Artificial Intelligence (AI) are increasingly becoming important in almost all healthcare processes [30–32]. The volume and diversity of technical assets present in healthcare institutions reflect the complexity of technology management, which must be effective to ensure that the equipment is always used safely and effectively. Investigating the factors that affect patient safety pushes health professionals to determine the causes of related problems, identify meliorative interventions, and evaluate the effectiveness and efficiency of such interventions. Patient safety is strictly related to health technologies, including devices, medicines, vaccines, procedures, and systems. Studying and managing health technology adverse events is critical for improving medical quality and safety [100]. In recent years, we have encountered a process of critical analysis of the manufacturer's maintenance recommendations, urging Clinical Engineers (CE) and Health Technology Management (HTM) professionals to adopt evidence-based methods to maintain medical equipment's dependability and safety while using their resources wisely [50].

In this scenario, Real-World Data (RWD), i.e., observational data associated with outcomes in real-world settings, can be used to generate Real-

World Evidence (RWE) to assess the effectiveness and the safety of a given health technology by examining the intended and unintended consequences of its use. RWE can be employed in healthcare for different purposes, such as to support more effective and cost-efficient medical product regulatory decision-making across the product life cycle. The new EU Medical Device Regulation 2017/745 (EU-MDR) requires companies to register their devices in the EUDAMED database following the European Medical Device Nomenclature (EMDN), and to provide a Periodic Safety Update report and a Post Market Surveillance (PMS) report [29]. Creating these reports strengthens the post-market monitoring and vigilance system of medical devices by improving quality and patient safety. The continuous analysis of the safety signals, which emerge from the adverse events of the medical devices available on the market, has indeed a strong significance for manufacturers in relation to the aforementioned legal obligations. Besides, RWE is also very useful for performing the market evaluation of a specific medical device, analysing faults, planning updates and interventions, and avoiding recalls. It is also well-recognized that RWE is a source for assessing the impact of health technologies in terms of risk minimisation, pricing, and reimbursement decisions [14].

Health Technology Assessment (HTA) is "a multidisciplinary process that uses explicit methods to determine the value of health technology at different points in its life-cycle. The purpose is to inform decision-making in order to promote an equitable, efficient, and high-quality health system" [54]. The process is formal, systematic, and transparent, and uses state-of-the-art methods and data collected during the routine delivery of health care to consider the best available evidence. Decision support techniques such as Multi-Criteria Decision Analysis (MCDA) or Analytic Hierarchy Process (AHP) are common methods employed in HTA which can be applied to obtain RWE on different aspects (e.g., safety and effectiveness) of different health technologies, in order to provide innovation in assessment for public international institutions and authorities, such as the World Health Organization (WHO) or National Health Systems (NHSs). Outcomes can be used to highlight both the most common faults and the unexpected new problems of medical devices.

Maintenance is another essential component of the activities in the hospital's CE and HTM departments, because of the enormous personnel and financial resources required. As a result, evaluating the efficiency of main-

tenance programmes solely depends on making the best use of the available resources [115]. Evidence-Based Maintenance (EBM) starts from the analysis of the causes of equipment failures and uses these results to continually improve maintenance. EBM involves the use of empirical data and scientific evidence to identify the optimal maintenance strategies for medical devices. This approach aims to improve the efficiency, reliability, and safety of medical equipment, which is critical to ensure the quality of healthcare delivery. EBM allows comparison of different maintenance strategies and provides concrete evidence to prove the safety and effectiveness of the one adopted. EBM begins with the analysis of RWE to monitor the maintenance effectiveness and plan any necessary changes to improve it. Maintenance reports, stored in Computerized Maintenance Management System (CMMS) software, can be a great source of RWD, from which RWE can be extracted. Unfortunately, CMMS often only contain a description of the failures, the repair procedures and any spare parts used, thus lacking information about any measures needed to prevent the failure [49].

Analyses based on RWE require the availability of a significant amount of RWD to perform a solid study and extract actual evidence of a general nature. Medical RWD can originate from different sources, such as Electronic Health Records (EHR), patient surveys, CMMS, and Spontaneous Reporting System (SRS) databases. One of the main difficulties which arises when dealing with RWD is the lack of standardization among countries which makes any sort of parsing nearly unfeasible. For instance, in the specific case of CMMS, the same medical device can be identified with different codes and different nomenclatures from country to country, as well as there is no international standard classification of fault codes.

Semantic ontologies are proven to be very useful tools which allow sharing as well as reusing concepts in a standardized way so that the data gathered from heterogeneous sources receive a common nomenclature [81]. They can be used to enhance the traditional approaches to healthcare facilities management, facilitate CE/HTM, HTA, strengthen communications, and outline every potential interaction between various roles.

Open-access SRS, such as the FDA Adverse Event Reporting System (FAERS), the Vaccine Adverse Event Reporting System (VAERS), the US Manufacturer and User Device Experience (MAUDE) database, and the EC EUDAMED are huge sources of information about adverse events related to health technologies, providing an enormous quantity of RWD which can

be freely accessed and further analysed to extract evidence. Recently, these data sources show the common trend of the gradual growth of adverse events related to Health Information Technologies (Fig. 1.1) [63], which is coherent with the diffusion of medical software in healthcare and with the resulting possible faults which, in addition, may be also caused by the hardware they are installed on [48].

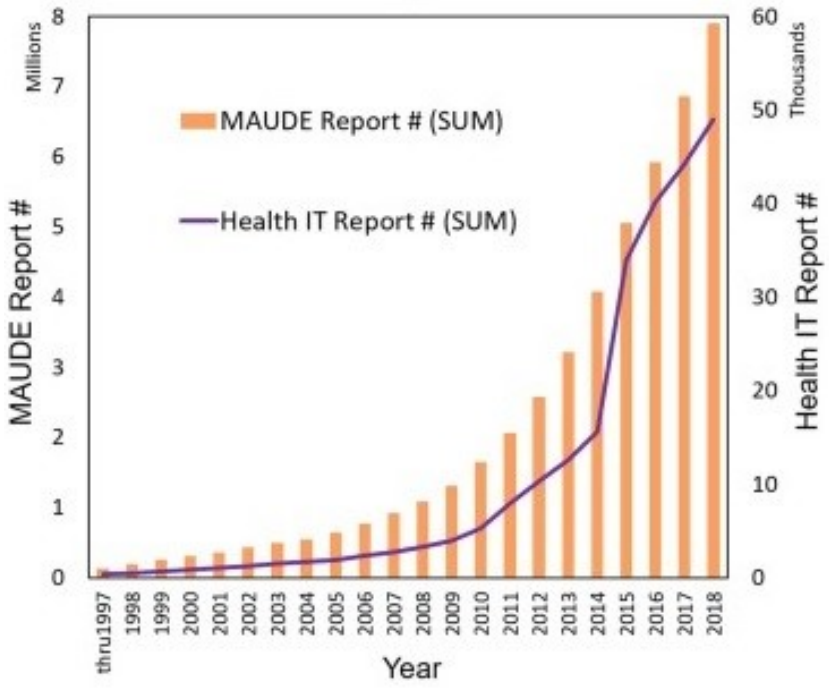


Figure 1.1: HIT-related adverse events extracted from the MAUDE database up to 2018 [63].

Health Information Technologies (HIT) are computing systems used in the storage, retrieval, analysis, and communication of health-related data. Internationally, billions of dollars have been invested in HIT and they are now routinely used to support the provision of healthcare, to increase the efficiency of health systems, and to improve the outcomes experienced by patients [83].

In such a scenario, using Neural Networks and Natural Language Processing (NLP) techniques to mine structured data from the above-cited sources can improve the extraction of RWE, thus empowering the processes of designing, assessing, evaluating, and managing health technologies. In such

real-world applications, explainability and transparency of AI systems are becoming more and more essential for users and for the researchers and developers who create AI solutions [123].

However, Neural Networks are usually weak in explaining their inference processes and final results, and they are typically treated as a black box by both developers and users. Explainable Artificial Intelligence (XAI) is a research field that aims to make AI systems results more understandable to humans. The term was first coined in 2004 by Van Lent et al. [103], to describe the ability of their system to explain the behaviour of AI-controlled entities in simulation game applications. Currently, the term XAI refers to the initiatives and efforts made in response to AI transparency and trust concerns, more than to a formal technical concept. The goal of enabling explainability in Machine Learning (ML) "is to ensure that algorithmic decisions, as well as any data driving those decisions, can be explained to end-users and other stakeholders in non-technical terms" [7].

1.1 The objective

The main goal of this work is to develop a framework to support the extraction of RWE through NLP and Deep Neural Networks (DNNs) for mining and classifying HIT adverse events extracted by heterogeneous sources of RWD. To achieve the proposed goal, records extracted by the MAUDE database have been labelled by experts as HIT/non-HIT adverse events and used to fine-tune a pre-trained model for binary text classification. The model has been validated with a 10-fold validation process and then tested against a subset of records to assess its performance. XAI methods have been further applied to highlight the most common features which led to a given classification, helping the final user understand the type of HIT-related adverse event. The developed framework is something that has not been experienced yet, and so it may result in a novel and possibly new successful approach for obtaining RWE for decision-making purposes in several Clinical Engineering fields, such as the above-cited Evidence-Based Maintenance, Health Technology Management and Assessment, and Post-Market Surveillance.

This work also intends to shed light on and elucidate the challenges arising from the noted absence of uniformity across nations, with a particular emphasis on the imperative for standardization and compatibility in the med-

ical device' nomenclature and faults classification. To accomplish this objective, a comprehensive literature scoping review was conducted on semantic ontologies to systematically chart research undertaken in ten designated domains of interest, with the ultimate goal of constructing a comprehensive and shared ontology that comprehensively depicts the healthcare environment.

Chapter 2

Background

2.1 Data sources available for medical device vigilance

Health authorities maintain two types of regulatory databases: Spontaneous Report System (SRS) databases and recall/alert databases. The main publicly available SRS databases are:

- The US Manufacturer and User Facility Device Experience (MAUDE), regulated by the Food & Drug Administration (FDA) Center for Devices & Radiological Health;
- The EU European Databank for Medical Devices (EUDAMED), regulated by the European Commission;
- The Australian Database of Adverse Event Notifications (DAEN), regulated by the Therapeutics Goods Administration.

Manufacturers are required to submit incidents to these vigilance databases. while healthcare professionals, patients, and other organizations may submit incidents at their discretion. Voluntary submissions are a leading cause of under-reporting. The FDA released the final ruling for medical device vigilance on a Unique Device Identification (UDI) system in 2013. The European Commission has also mandated UDI adoption since 2020 for devices obtaining certification under the EU-MDR. The International Medical Device Regulators Forum (IMDRF) created a dictionary of adverse events, device

malfunction, and investigation codes specific to medical devices, which is being implemented both in the FDA MAUDE and in the EC EUDAMED [53]. MAUDE is the most widely used and publicly accessible SRS, collecting data from all around the world [16]. EUDAMED is expected to have both UDI and IMDRF coding for devices and events, but event submissions have only started in 2022, while public database access levels still remain uncertain.

Table 2.1: Publicly Available Vigilance Databases.

Country	**National Regulatory Authority**	**Database**
United States	FDA Center for Devices & Radiological Health	Manufacturer and User Facility Device Experience (MAUDE)
European Union	European Commission	European Databank for Medical Devices (EUDAMED)
Australia	Therapeutics Goods Administration	Database of Adverse Event Notifications (DAEN)

2.2 Spontaneous Reporting Systems and Health Information Technologies fault classification

A literature review has been performed on SRS databases and their use for data and text mining, as well as on Health Information Technology (HIT) fault classification. The majority of works in the area of mining SRS databases were based on the FDA Adverse Event Reporting System (FAERS) and the Vaccine Adverse Event Reporting System (VAERS) but also included articles based on the MAUDE. Very few of them regarded data repositories maintained by other countries. Identified articles focused only on structured data. However, some investigators have recently begun to tune the SRS disproportionality results with information from other data repositories. For example, Harpaz et al. [42] and Iyer et al. [60] utilized information from the clinical notes of electronic health records to augment

the results of signal detection in the FAERS database, while Xu et al. [124] used the biomedical literature within MEDLINE to boost the results of disproportionality analysis. A study by Wang et al. [117] focused on creating a normalized, open-source data mining set of FAERS drug information aggregated with RxNorm, the National Drug File-Reference Terminology and the Medical Dictionary for Regulatory Activities (MedDRA). This work should facilitate downstream text mining approaches within FAERS. It also highlights the level of detailed effort needed to collate the information in SRS databases such as FAERS, VAERS, and MAUDE, where entry of non-structured product identifying data is the norm.

A brief literature review has also been performed on adverse events solely referred to HIT. Alemzadeh et al. [2] studied 5,294 medical device (MD) recalls between 2006 and 2011 and observed that computer-related recalls contributed to 1,210 of all recalls in the period. Of the computer-related recalls, 94% presented some risk of serious injury or death. 64% of the computer-related recalls were related to software faults. Studies on HIT failure [73] show that from January 2015 to July 2017, there were 678 reports of 436 different adverse medical device events associated with health information technology. Most of the 46 events associated with patient harm were related to the computerized physician order entry and picture archiving and communication systems. Software issues were classified into four categories: functionality, system configuration, device interface, and network configuration. An analysis of MD recalls registered in FDA records for the period 1999-2005 reported that one-third concerned MDs using software for their functioning and showed a constant increase of software failure throughout these years [11].

2.3 Natural Language Processing

A literature review has been performed on Natural Language Processing (NLP) techniques applied to medical devices. Sentiment Analysis (SA) is an approach to NLP that identifies the emotional tone behind a body of text. In healthcare, it can be employed to gain insights from both medical social media and clinical documents regarding the effectiveness of a treatment or medication [19].

A recently published scoping review was used as a starting point [89]. The first distinction that emerges from the review is the preferred approach

used for NLP tasks and SA: lexicon-based vs ML-based. As generally stated by the authors, the lexicon-based approach is not well suited for capturing the meanings in medical texts. Comparing the outcomes of studies which used both lexicon and ML approach [61] with the outcomes of studies which used the first [71] or the second method [23] (same topic and data sources) confirmed the findings from previous works: the lexicon-based approach has significant drawbacks in evaluating the real sentiment of health technologies. The majority of analysed papers rely on ML methods (Support Vector Machines - SVM, Naive-Bayes, regression tree) using input features such as POS (Part of Speech) tagging, TD-IDF (Term Frequency-Inverse Document Frequency), BTO (Binary Term Occurrences) and Word2Vec. In particular, the SVM classifier is one of the most successfully used in opinion mining [23,61].

Additional domain-specific features have also been explored in some approaches, mainly UMLS (Unified Medical Language System) concepts reflecting medical conditions and treatments, such as MeSH [21]. Other works, focusing on NLP but outside the medical domain, also used external resources for query expansion such as DBpedia [84, 85] or Babelnet [75]. Biomedical texts, including adverse events reports, are potential resources of massive information and hidden knowledge, unfortunately, it is not possible for researchers and practitioners to keep themselves updated with all the developments in the biomedical field [80]. The emphasis of biomedical research is therefore shifting from individual entities to whole systems, with the demand of extracting relationships between entities from biomedical text to generate knowledge.

Biomedical Causal Relation Extraction (BCRE) aims to efficiently reveal high-quality relations from domain-related resources [126]. Recent works have shown how deep learning can be used to solve NLP and BCRE tasks. Deep-learning models exemplifying the notion of unsupervised representation learning are the autoencoders (AE). They became popular as an early tool to pre-train supervised deep-learning models, especially when labelled data was scarce, but still retain usefulness for entirely unsupervised tasks such as phenotype discovery [96]. AEs have been applied to biomedical fields for clinical relation extraction [72] and, in EHRs, as outcome predictors [78].

Transformer models, introduced in 2017 [104], have been developed to address NLP tasks and overcome the limitations of Recurrent Neural Networks (RNNs). Transformers have become the deep-learning model of choice for NLP problems [65, 119] (Fig. 2.1).

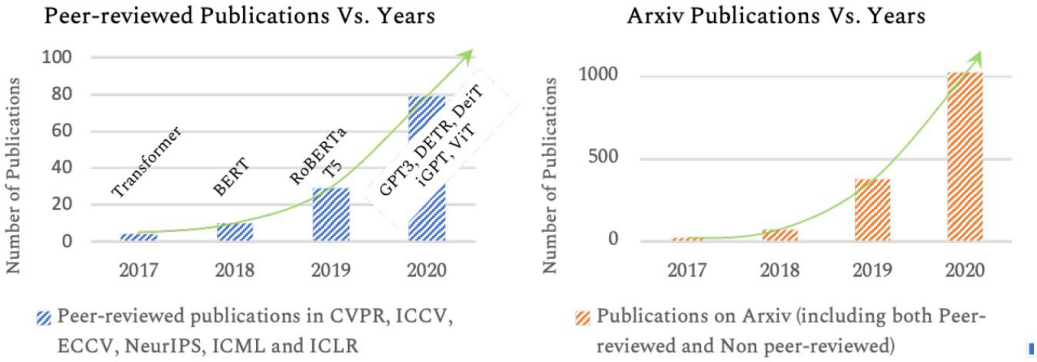


Figure 2.1: Peer-reviewed and non-peer-reviewed publications about NLP over the years [65].

Transformers make use of multi-headed self-attention to perform sequence-to-sequence learning tasks [104]. Self-attention is used to learn long-range dependencies between the elements in a sequence. Multi-head self-attention is the combination of several attention heads. This is conceptually similar to how a convolutional layer can consist of multiple convolution filters, with each filter independently extracting different types of features. The attention mechanism performs a lookup producing a set of weights for each element. The most relevant elements have the highest attention scores. This allows the model to be explainable with reference to both input and output.

This has led to the development of pre-trained systems such as the popular BERT (Bidirectional Encoder Representations from Transformers) [20] and GPT (Generative Pre-trained Transformer) [91]. They are trained with huge general-language datasets such as Wikipedia Corpus and Common Crawl. They can be also fine-tuned to specific NLP tasks [127] (Fig. 2.2).

NLP is particularly booming in the healthcare industry. This technology is improving care delivery and disease diagnosis, as well as medical equipment management while bringing costs down while healthcare organizations are going through a growing adoption of electronic data management systems (Fig. 2.3).

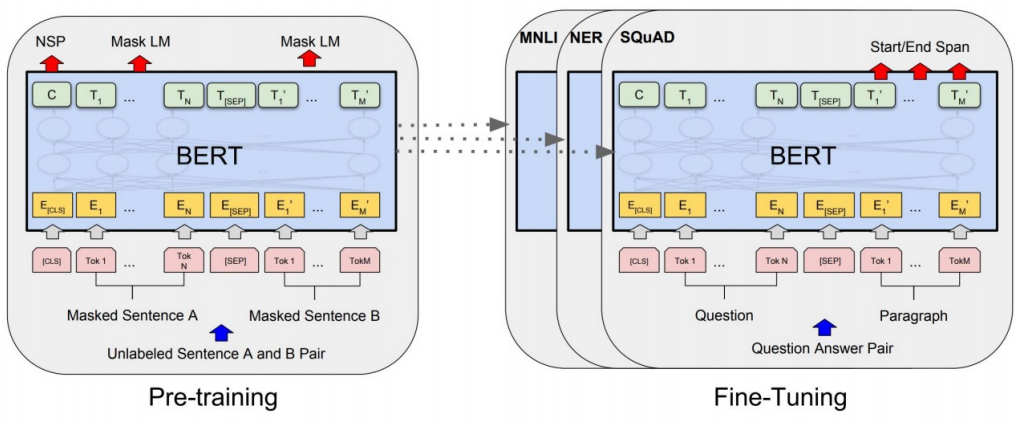


Figure 2.2: BERT acts as an effective pre-training for several NLP tasks. The model can be simply fine-tuned on specific tasks by changing the classification head [20].

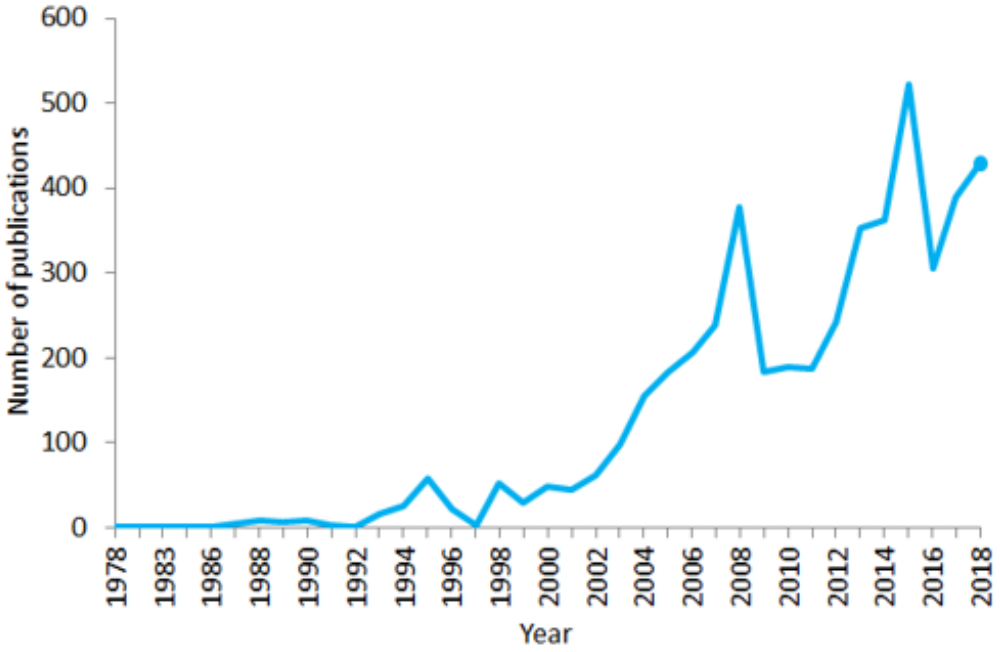


Figure 2.3: Number of publications containing the sentence "natural language processing" in PubMed in the period 1978-2018. As of 2018, PubMed comprised more than 29 million citations for biomedical literature.

Biomedical and clinical NLP community researchers have been actively

proposing BERT-based models to process biomedical and clinical text information effectively and efficiently. Successful biomedical and clinical BERT models include SciBERT [10], which is built on basic BERT to increase its performance on scientific data; BioBERT [67], which is a domain-specific BERT model pre-trained on a large number of biomedical text corpora; ClinicalBERT [46], which is a language representation model to extract high-quality relationships between medical concepts from extensive clinical notes. Recent studies [1] show that these specifically pre-trained models outperform classical models in extracting evidence from biomedical corpora.

2.4 Semantic Ontologies

A semantic ontology is a tool for knowledge representation built on formal collections of terms. It is used to describe and represent a field of interest (also known as a domain) clearly and consistently. The Semantic Web (Fig. 2.4), a World Wide Web extension created by the World Wide Web Consortium (W3C) [109] with the primary objective of enabling computers to support networked interactions, is grounded on ontologies. The Semantic Web offers a framework for data querying and ontology-based inferences using a variety of technologies. Numerous applications use ontologies and vocabularies to make it easier to integrate data from various sources and to formally organize knowledge by connecting terms through logical relationships. Drawing inferences (automatic processes that create new relationships based on the data stored in the vocabulary itself) in order to carry out reasoning procedures is also made possible by ontologies. An ontology's structure is hierarchical and is based on techniques that divide the items it contains into "classes" and "sub-classes". Individual resources may then be mutually associated, resulting in the logical association of classes and instances. Semantic ontologies are becoming increasingly important because of their capabilities to provide a common representation of a domain among different users by linking concepts and instances, supporting interoperability between heterogeneous data archives, and fostering the reuse and sharing of knowledge [64].

The W3C provides several techniques to define various forms of standard vocabularies given the broad range of operations provided by ontologies, such as Resource Description Framework (RDF), Web Ontology Language (OWL), Javascript Object Notation for Linked Data (JSON-LD), and HL7

Fast Healthcare Interoperability Resources (FHIR) [107]. According to the W3C Semantic Web, RDF is a standard model for data interchange on the Web. RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed. RDF extends the linking structure of the Web to use Uniform Resource Identifiers (URIs) to name the relationship between things as well as the two ends of the link (this is usually referred to as a "triple"). Using this simple model, it allows structured and semi-structured data to be mixed, exposed, and shared across different applications. This linking structure forms a directed, labelled graph, where the edges define the link between two resources, represented by the graph nodes. This graph view is the easiest mental model for RDF and is often used in easy-to-understand visual explanations [106]. OWL is a language for the semantic web, designed to represent rich and complex knowledge about things, groups of things, and relations between things. OWL is a computational logic-based language making it possible that knowledge expressed in OWL can be exploited by computer programs, for verifying consistency or to make explicit some implicit knowledge. OWL documents, known as ontologies, can be published on the World Wide Web and may refer to or be referred by other OWL ontologies [105]. JSON-LD is a linked data serialization recommended by the W3C. It is an extension of the JSON format that integrates Linked Data to a website. It also provides an RDF serialization format to contextualize data [108]. HL7/FHIR is a standard for exchanging electronic healthcare information allowing data requests and transfers between various healthcare systems. The main goal of FHIR is to solve a wide range of clinical and administrative healthcare problems to improve interoperability; it can be expressed as XML (eXtensible Markup Language), JSON, or RDF/TURTLE encodings [45].

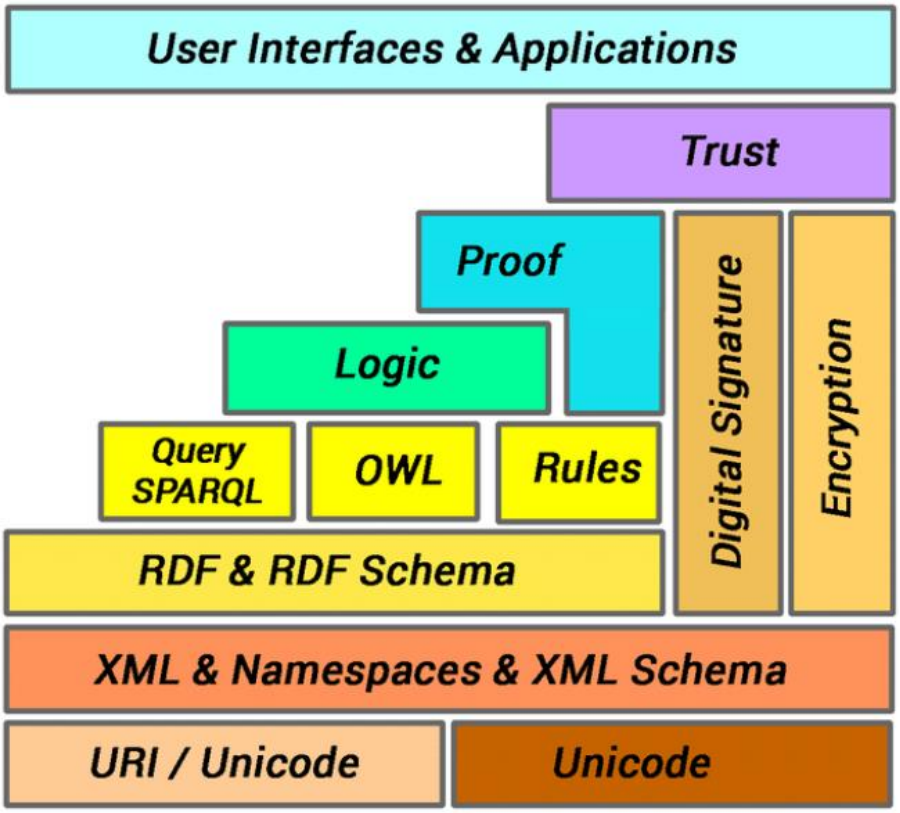


Figure 2.4: Semantic Web stack.

Chapter 3

Semantic Ontologies for Evidence-Based Maintenance in Complex Healthcare Structures

In this chapter, the challenges that can arise in the realm of Evidence-Based Maintenance (EBM), particularly when confronted with the absence of standardised naming and coding conventions for medical devices, are highlighted and articulated. The emphasis is placed on the nomenclature of medical equipment and the standardisation of fault codes, and how a global harmonised nomenclature can overcome the lack of global standards, which contributes to the scarcity of accessible and shareable data to extract evidence.

In this regard, semantic ontologies have the potential to establish a level of abstraction for standardised concept sharing and reuse. By doing so, it is ensured that data from many sources can be provided with a standard nomenclature, improving stakeholder communication. A scoping review was conducted to identify and examine existing ontologies capable of encompassing the heterogeneity of technologies that are currently associated with the hospital environment. The review was carried out on the Scopus database on January 13th, 2023, utilizing the PRISMA extensions designed for scoping reviews. A total of 3,225 documents resulting from the database search were screened by two reviewers. Subsequent refinement led to a final selection of 32 articles for in-depth analysis. Furthermore, a total of 34 ontologies extracted from the identified articles were subjected to analysis and discussion. The outcomes of this study are anticipated to pave the way for the

development of the ODIN Ontology and the OdinEMDN Ontology within the EU Project ODIN (see Section 3.1.3). These unified integrated ontologies are envisioned to encompass information about healthcare entities and their semantic relationships, thereby fostering enhanced data exchange and interconnectivity among individuals, devices, and applications within an expanded framework that includes the Internet of Things (IoT), robotics, and Artificial Intelligence (AI).

3.1 Introduction

3.1.1 Nomenclature of Medical Devices

The nomenclature of medical devices is a coding and naming system used to classify and identify all medical devices and related health products. According to different classification and nomenclature systems, 5,000 to 24,000 different types of medical devices can be identified, ranging from very simple to complex, inexpensive to costly. Fig. 3.1 clearly shows the heterogeneity of the existing nomenclature systems, highlighting that 39% of countries do not use any nomenclature, 8% use more than one system, and 16% have a nationally developed one.

The nomenclature systems most widely used for medical devices are the European Medical Device Nomenclature (EMDN), the Global Medical Devices Nomenclature System (GMDN), and the Universal Medical Devices Nomenclature System (UMDNS).

The **EMDN** is the nomenclature of use by manufacturers when registering their medical devices in the EUDAMED database according to the EU Medical Devices Regulation 2017/745 [29]. Founded on pre-established criteria and requirements and based on orientations provided by the Medical Device Coordination Group (MDCG), the European Commission decided in favour of the use of the "Classificazione Nazionale Dispositivi medici" (National Classification of Medical Devices - CND) as the basis for the EMDN.

The **GMDN** was developed by the European Committee for Standardization (CEN) and medical device experts from around the world (manufacturers, healthcare authorities and regulators) based on the international standard ISO 15225 [57]. It is managed and maintained by a not-for-profit company, the GMDN Agency, which reports to a Board of Trustees on which medical device regulators and industry are represented. To ensure the con-

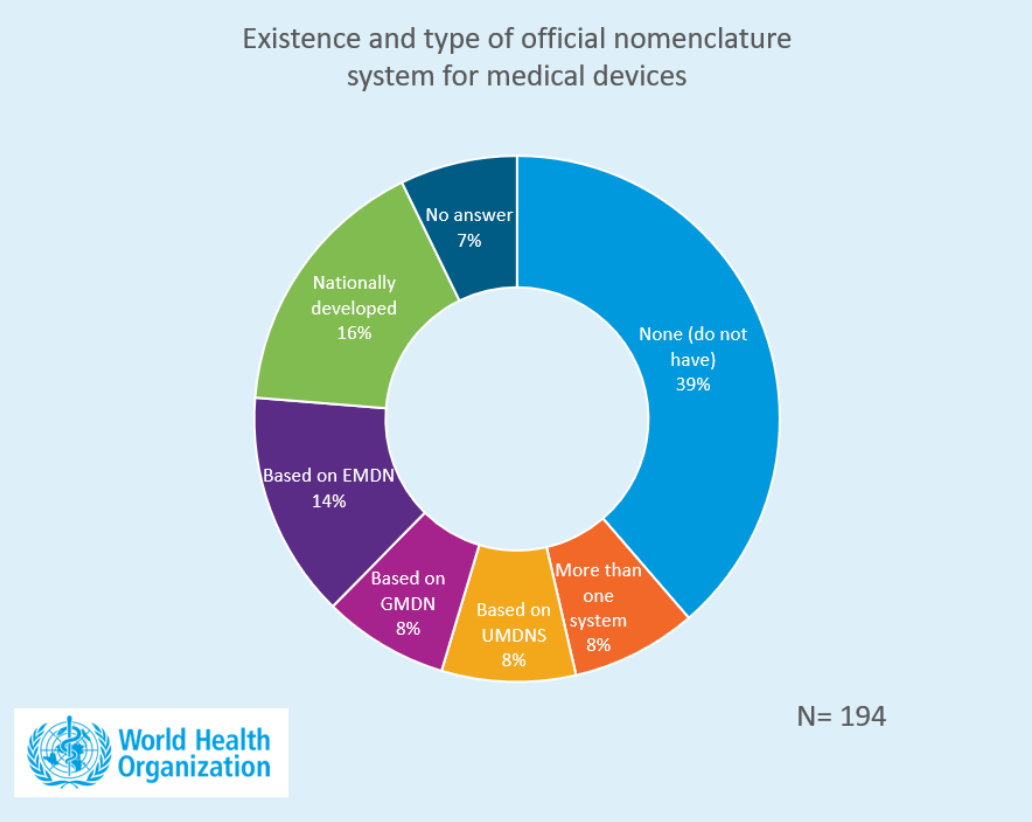


Figure 3.1: Distribution of countries based on the implemented nomenclature system [122].

tinuing permanency of the GMDN, revenues are generated through the licensing and sale of GMDN Agency products, particularly the GMDN codes. The GMDN is a poly-hierarchical system. Product identification is done by unique numerical five-digit numbers that are associated with a term (medical device name), a definition that includes the intended use(s) and the device category (based on device application, technology, or other common characteristics). Identification of all specific medical devices having substantially similar generic features is possible through cross-referencing

The **UMDNS** was developed by the Emergency Care Research Institute (ECRI). ECRI is a nongovernmental, not-for-profit organization, governed by an Executive Committee and a Board of Trustees. The UMDNS is poly-hierarchical and is developed as an interrelated vocabulary based on terms naming the medical devices. Terms are assigned a 5-digit code using consecutive numbering with no intrinsic meaning. The code is associated with

a definition and a description of the intended use. Associated properties provide additional attributes for classification.

The multiple nomenclatures in existence make it difficult to communicate important information between individuals and organisations, which can result in negative health, economic, and social impacts. It complicates interoperability, data extraction, procurement, supply and trade, and tracking of medical devices, negatively affecting patient safety, as well as technology management and maintenance [47]. Having a nomenclature system in place for medical equipment would facilitate its management and regulation by standardizing terms that enable communication despite linguistic and other barriers. Such standardisation should be a prerequisite for inventory management and databases for the maintenance of equipment since it would provide a globally accessible, transparent and harmonized nomenclature system. The World Health Organization (WHO) is one of the most significant international entities involved in the effort to establish a universal nomenclature for medical devices. During the 145th WHO Executive Board in 2019, the Director-General emphasized the necessity for a standardised nomenclature of medical devices "as a common language for recording and reporting medical devices across the entire health system at all levels of health care for a full range of purposes [...] The lack of a nomenclature system has hampered the development of the evidence and web-based health technologies database to provide guidance on appropriate medical devices" [120]. Besides, such a lack is actually impeding progress towards access to medical devices, which has a negative impact on efforts to facilitate emergency interventions and achieve universal health coverage [122].

WHO recognizes the availability of multiple systems and offers a platform towards convergence (Fig. 3.2).

WHO presented the first development of the International Classification and Nomenclature of Medical Devices (ICMD), implemented in the ICD-11 (International Classification of Diseases) platform. The classification and terms generated represent the harmonisation of nomenclature and classes in the form of ontology and it is still under development. During the last 152nd WHO Executive Board in 2022, the Director-General still focused the attention on the fact that "the goal is to create a standardized international classification, coding and nomenclature for medical devices that would be available to all Member States and would support patient safety, access to medical devices for universal health coverage, emergency preparedness and

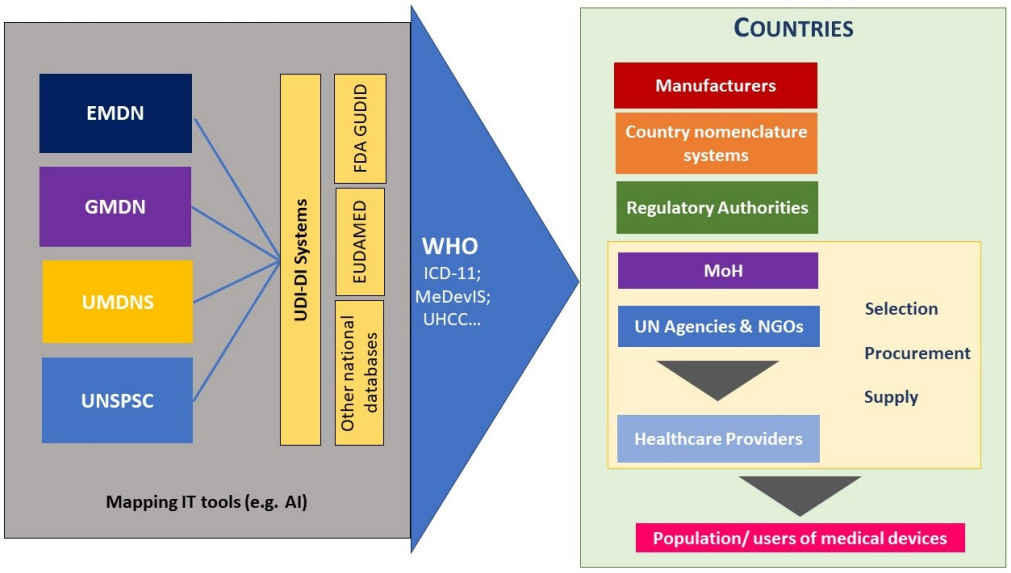


Figure 3.2: Graphical overview of the process of mapping across different nomenclature systems leading to the implementation of the ICD [122].

response, efforts to increase the quality of health care". Moreover, a request for proposals was posted in the United Nations Global Market Place from September 26th to October 20th 2022, with the intention of entering into a contract with the successful bidder for the provision of mapping medical device nomenclature data for integration in WHO platforms. The goals of the request were:

- 1,200 types of medical devices due 3 months.
- 4,000 types in 5 months.

Contractors from four different countries were among the six offers that were examined. The request will be fulfilled by Symmetric Health Solutions from December 2022 to July 2023 [121].

3.1.2 Standardization of Failure Codes for maintenance

Computerized Maintenance Management System (CMMS) software became essential for Health Technology Management (HTM) program operations. In fact, modern CMMS software contains the fundamental fields that are needed for basic HTM. Unfortunately, HTM programs differ widely in how they

configure or suggest the use of those fields. More broadly, the lack of standardisation largely prevents benchmarking between different structures and implementations. Performance metrics from one HTM program often cannot be compared to metrics from another HTM program. As a consequence, the HTM community stays in a weak position when confronting regulatory and accreditation agencies. In response to this challenge, the Association for the Advancement of Medical Instrumentation (AAMI) sponsored a "CMMS Collaborative" project among CMMS suppliers. The project started with the assumption that better use of existing CMMS software would make it easier to feed databases with accurate data and extract useful information from it. The involved suppliers all agreed on proposing a standard for the "Failure Code" field, as its purpose is to document the reason why a piece of medical equipment was unable to achieve its clinical objective of diagnosis, treatment, or monitoring [6]. Table 3.1 shows the proposed failure codes.

Table 3.1: Failure Code field options proposed by AAMI's CMMS Collaborative project [6].

Code	**Description**
Accessory or Disposable Failure	Failure of device accessory or disposable, not a failure of the device itself.
Calibration Failure	Failure of a device to meet calibration parameters, requiring recalibration.
Component Failure (Battery)	Failure of the battery that provides power for device operation.
Component Failure (Not Battery)	Failure of a device component other than the battery.
Failure Caused by Maintenance	Failure of a device resulting from maintenance activities.
Failure Caused by Abuse or Negligence	Failure of a device resulting from damage caused by intentional misuse or negligent use.
Network or Connectivity Failure	Functional failure external to device from failure of network or connectivity.
Software Failure	Functional failure of a device resulting from malfunctioning software.
Use Error (Use Failure)	Failure of a device to support achievement of a clinical objective.
Failure Caused by Utility System	Functional failure of a device resulting from failure of or access to a utility system.
Failure Cause by Environmental Factor	Functional failure of a device resulting from an environmental factor.
Failure Could Not Be Identified	Reported failure could not be reproduced or identified by testing.
Failure Not Diagnosed-Device Not Repaired	Reported failure indicated that testing or repair was unwarranted.
No Failure Associated with the Work Orders	There was no failure associated with the work order (included for completeness).

A previous study by Iadanza et al. [49] showing the application of the EBM approach to a large hospital fleet of electromedical equipment, proposed a set of fault codes for corrective and predictive maintenance (Table 3.2), derived by [111–114], with the final goal of calculating 20 Key Performance Indicators (KPIs).

Table 3.2: Failure codes for corrective and predictive maintenance proposed by [49].

Code	**Description**
NPF	No problem found
BATT	Battery failure
ACC	Accessory failure (including supplies)
NET	Failure related to network
USE	Failure induced by use (i.e., abuse, accident, environment conditions)
UPF	Unpreventable failure caused by normal wear and tear
PPF	Predictable and preventable failure
SIF	Induced by service (i.e., caused by a technical intervention not properly completed or premature failures of a part just replaced)
EF	Evident failure (i.e., evident to the user but not reported)
PF	Potential failure (i.e., in the process of occurring)
HF	Hidden failure (i.e., not detectable by the user unless special test or measurement equipment)

By analysing the mentioned examples, it surely emerges that the process toward standardisation of CMMS failure codes has already begun. On the other hand, a lack of consistency is still missing: the mentioned project by AAMI has been performed with limited consideration of the existing academic literature. A sore point is that the proposed fault codes are mostly focused on hardware failures, leaving it open to the challenge of understanding the trends in the faults related to Health Information Technologies (HIT), considering that medical software is becoming more and more pervasive in healthcare.

3.1.3 The ODIN Project

ODIN is a European project funded under Horizon 2020 [28] - the EU Research and Innovation program that has the aim to achieve the generation of world-class science - dedicated to advancing hospital safety, productivity, and quality. The primary objective of the project revolves around the development and delivery of an open digital platform, bolstered by the integration of robotics, IoT solutions, and specialized artificial intelligence, aimed at providing a comprehensive suite of services and Key Enabling Resources (KERs). These resources are rigorously tested in the healthcare environments of leading hospitals across six European nations: Spain, France, Germany, Poland, the Netherlands, and Italy [86]. The project's implementation focuses on three distinct areas within the hospital setting, referred to as eWorkers, eRobots, and eLocations. The eWorkers emphasizes equipping hospital staff with technology solutions to alleviate their burdensome and time-consuming tasks while enhancing routine activities. The eRobots facet concentrates on automating hospital processes through the deployment of robotic technology, thus assisting human workers in their roles and enabling them to focus on core responsibilities. Finally, eLocations aims to make medical facilities smarter by deploying suitable instrumentation. Medical locations are equipped with sensors, technologies that facilitate human interaction, and high connectivity to effectively communicate with hospital staff, robots, and devices. These intervention areas collectively address a wide array of crucial aspects within the hospital environment, spanning logistics, robotics, IoT, and disaster management.

3.2 Methods

3.2.1 Information Source

The literature search was carried out through the Scopus database[1] on January 13th 2023 using the PRISMA extensions for scoping reviews [101]. The initial search results were screened by two different reviewers (CP and AL) using a selection based on titles. A further evaluation was performed by the same reviewers on the basis of the abstracts for the selected results. At this stage, a third reviewer (EI) ruled on possible inconsistencies. All three re-

[1] https://www.elsevier.com/solutions/scopus

viewers were involved in the final selection of full texts for potentially relevant publications. At this stage disagreements on study selection were resolved by discussion among the reviewers.

3.2.2 Search

Carefully selected keywords were given as input to the Scopus search engine. They were selected according to the specific scope of the review to find and select ontologies consistent with the selected areas of intervention (see Section 1.1). Besides, the chosen keywords should also reflect the aspects linked to clinical engineering, logistics, and disaster preparedness more adequately. The selected keywords are the following: IoT, IoT Healthcare, Drugs Robotics, Emergency, Disaster, Clinical Workflow, Surgery, Logistics, Data Collection, Staff, and Medical Record. All the above followed by the words "semantic ontology".

3.2.3 Eligibility and Exclusion Criteria

The majority of the articles that were targeted for the research addressed the subjects mentioned in section 1, including robotics, the Internet of Things, healthcare and the hospital environment, logistics management, medical personnel, data collection, and disaster preparedness and management. The search was restricted to documents produced after the year 2000 (included), written in English. Only scientific articles and reviews were included, leaving outside all other academic publications and all materials and research produced by organisations outside of the academic publishing (grey literature) to provide a high level of reliability and integrity. The included subjects are those which the author thought to be consistent with the related ten areas of intervention (see Section 1.1): computer science, engineering, medicine, social sciences, decision sciences, multidisciplinary, business, management and accounting, health professions, environmental science, pharmacology toxicology, and nursing. The exclusion criteria were created to prevent the selection of articles that discussed ontologies that were not publicly accessible or that belonged to a domain unrelated to the project's goals. The final Scopus database query is:

```
TITLE-ABS-KEY (semantic* AND ontolog*) AND TITLE-ABS-KEY ("IoT" OR
"Health" OR "Healthcare" OR "Robot*" OR "Emergenc*" OR "Disaster*" OR
```

```
"Clinic*" OR "Workflow*" OR "Surger*" OR "Logistic*" OR "Data*" OR
"DATA AND Collect*" OR "Staff" OR "Medical record*" OR "Internet AND
of AND things") AND (LIMIT-TO (OA,"all")) AND (LIMIT-TO
(PUBSTAGE,"final")) AND (LIMIT-TO ( DOCTYPE,"ar") OR LIMIT-TO
(DOCTYPE,"re")) AND (LIMIT-TO (SUBJAREA,"COMP") OR LIMIT-TO
(SUBJAREA,"ENGI") OR LIMIT-TO (SUBJAREA,"MEDI") OR LIMIT-TO
(SUBJAREA,"DECI") OR LIMIT-TO (SUBJAREA,"SOCI") OR LIMIT-TO
(SUBJAREA,"BUSI") OR LIMIT-TO (SUBJAREA,"ENVI") OR LIMIT-TO
(SUBJAREA,"MULT") OR LIMIT-TO (SUBJAREA,"HEAL") OR LIMIT-TO
(SUBJAREA,"PHAR") OR LIMIT-TO (SUBJAREA,"NURS")) AND (LIMIT-TO
(PUBYEAR,2023) OR LIMIT-TO (PUBYEAR,2022) OR LIMIT-TO (PUBYEAR,2021)
OR LIMIT-TO (PUBYEAR,2020) OR LIMIT-TO (PUBYEAR,2019) OR LIMIT-TO
(PUBYEAR,2018) OR LIMIT-TO (PUBYEAR,2017) OR LIMIT-TO (PUBYEAR,2016)
OR LIMIT-TO (PUBYEAR,2015) OR LIMIT-TO (PUBYEAR,2014) OR LIMIT-TO
(PUBYEAR,2013) OR LIMIT-TO (PUBYEAR,2012) OR LIMIT-TO (PUBYEAR,2011)
OR LIMIT-TO (PUBYEAR,2010) OR LIMIT-TO (PUBYEAR,2009) OR LIMIT-TO
(PUBYEAR,2008) OR LIMIT-TO (PUBYEAR,2007) OR LIMIT-TO (PUBYEAR,2006)
OR LIMIT-TO (PUBYEAR,2005) OR LIMIT-TO (PUBYEAR,2004) OR LIMIT-TO
(PUBYEAR,2003) OR LIMIT-TO (PUBYEAR,2002) OR LIMIT-TO (PUBYEAR,2001)
OR LIMIT-TO (PUBYEAR,2000)) AND (LIMIT-TO (LANGUAGE,"English"))
```

3.3 Results

The literature search led to a total of 3,225 articles, hence a selection was performed by two reviewers. The flow diagram in Fig. 3.3 illustrates the procedure for choosing the literature that was included in the final review.

714 documents out of the records obtained from the initial search were considered relevant after reading the title. The abstracts of the articles belonging to this set of items were then analysed for an additional screening, which resulted in the selection of 183 items by both reviewers. A third reviewer ruled on the 18 discordant opinions and selected 6 publications for a total of 189 selected records. Finally, 32 documents have been selected by the reviewers after reading the full text of the remaining articles. All the articles included in the review have been produced between the years 2010 and 2023. The main characteristics of the documents that were selected are displayed in Table A.1. The columns of such table are arranged in the following order: the first column shows the first author mentioned in the article, the year of

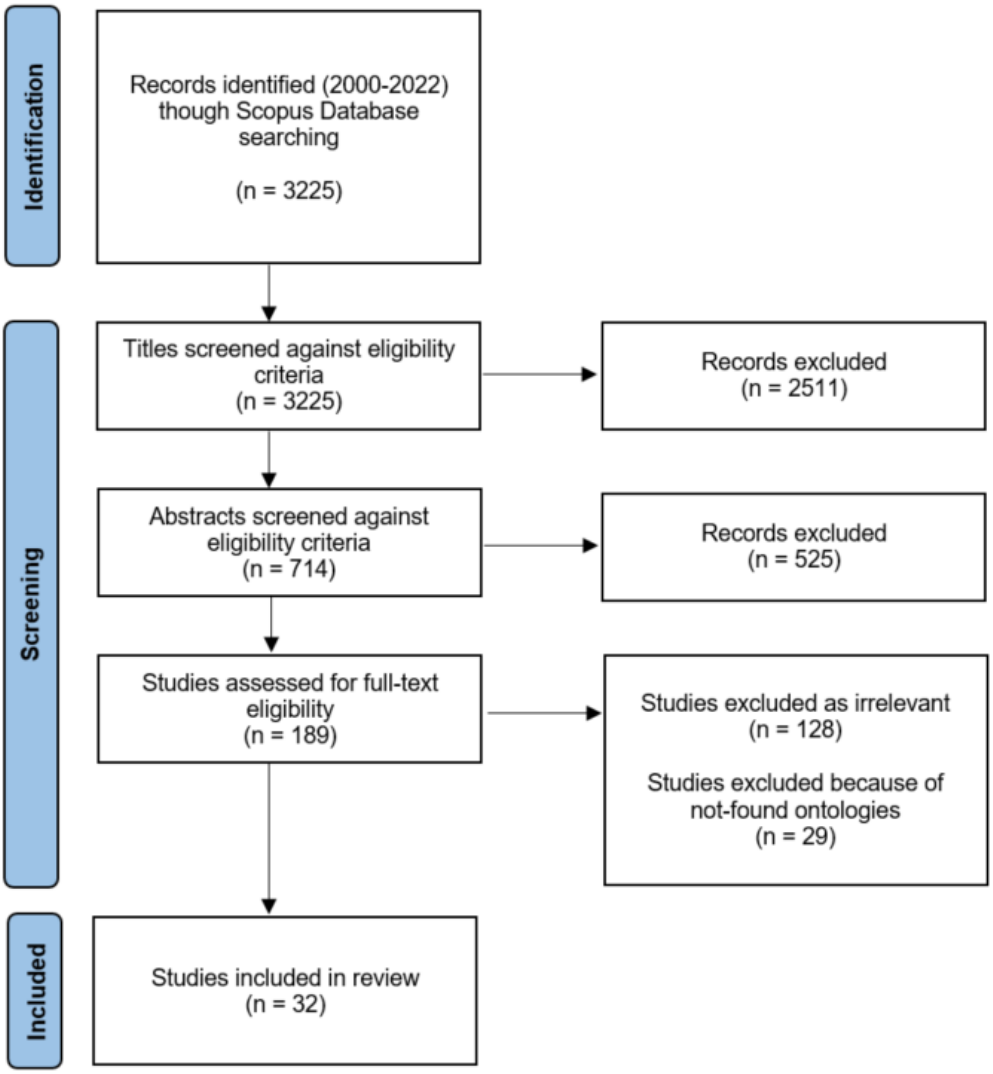


Figure 3.3: Flow diagram representing the process of selection of the included studies.

publication and the bibliography reference, the subsequent columns indicate, respectively, the first author and the year of publication, the title, the aim of the article under discussion, the mentioned ontologies, and the semantic domain that serves as the framework for the document's coverage.

3.3.1 Selection of sources of evidence for ontologies

The analysis of the results aims to identify public ontologies which can provide a semantic representation of the aforementioned topics and needs. Two main databases have been used to identify suitable public biomedical ontologies: Ontobee [34] and BioPortal [82].

Ontobee is a linked data server designed for ontologies that provides the query, visualization and comparison of different ontologies and ontology terms. It represents the default server for biomedical ontologies in the Open Biological Ontology (OBO) Foundry, a group of researchers that aim to establish a set of principles to follow when developing ontologies for the biological sciences. Basic Formal Ontology (BFO) is the official top-level ontology for all OBO Foundry ontologies. BFO is frequently used as ontology top-level architecture [4] and has been approved as international standard ISO/IEC 21838-2 [58].

BioPortal is an open repository of biomedical ontologies delivered by the National Centre for Biomedical Ontology (NCBO), which was formed as part of the National Centers for Biomedical Computing network founded by the National Institutes of Health (NIH). The goal of NCBO is to support biomedical researchers by providing online tools such as BioPortal, which contains ontologies concerning anatomy, chemistry and health.

3.3.2 Synthesis of results

A total of 34 ontologies, extracted from the selected articles, have been collected and reviewed. All of them are either represented in OWL or RDF format, according to W3C standards and are all accessible online for browsing and downloading. Table A.2 displays the applicable ontologies, a brief description of the represented domain, the Internationalized Resource Identifier (IRI), the source which they can be downloaded from, the main topic, and the referenced article.

3.3.3 The ODIN Ontology and the OdinEMDN Ontology

The introduction of technologies within the ODIN Project is geared toward enabling the real-time management of medical devices. This is made possible through the combined use of AI, IoT, robots, sensors, wearable devices

for staff, and a semantic web architecture solution. The final objective is to encourage greater alignment and standardization in Health Technology Management across European hospitals. In this context, the primary goal is to address the challenge of inadequate real-time information exchange among hospital personnel. At a semantic level, all these objectives are achieved through the development of the ODIN Ontology (Fig. 3.4), an ontology capable of defining various hospital structures, allowing for the reusability of this ontology. Many of the classes, properties, and individuals of the ontologies identified by the scoping review have been used to construct the ODIN Ontology:

- NCIT Ontology includes prefixes **thesaurus:** and **obo:**
- SCTO includes the prefix **snomed:**
- CORA Ontology includes prefixes **sumo-cora:** and **cora-bare:**
- BOT Ontology includes prefixes **bot:** and **building:**
- WoT Ontology includes prefixes **wot:**, **om:**, and **core:**
- ICD9CM Ontology includes the prefix **icd9cm:**
- The Organizational Ontology includes the prefix **org:**

Medical devices play a crucial role in the implementation of the ODIN Ontology. Explicitly pinpointing the precise location of medical devices, identifying their users, and tracking their usage conditions are essential for achieving a seamless real-time information flow. Consequently, the development of an ontology that comprehensively encompasses all existing medical equipment becomes a necessity to ensure their future utilization. In light of the absence of such an ontology, the European Medical Device Nomenclature Semantic Ontology (OdinEMDN) was created by leveraging the EMDN, filling the critical gap in this regard. This newly developed ontology became an integral component of the ODIN Ontology itself, including the prefix **odinemdn:**.

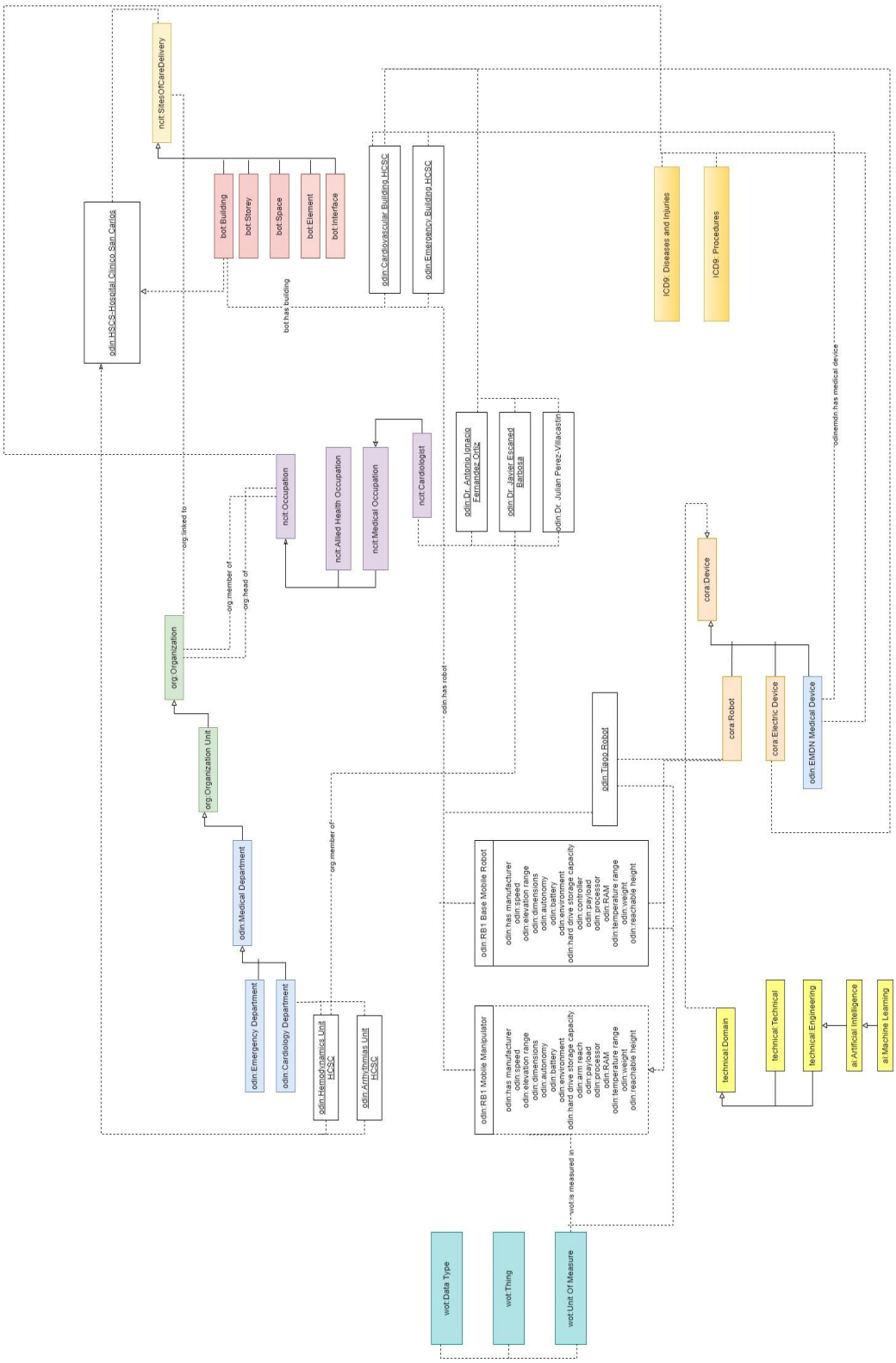


Figure 3.4: The ODIN Ontology.

3.4 Discussion

As displayed in Table A.1, each of the studies that were chosen is focused on a particular ontology and has been connected to a semantic domain in line with the review's objectives. Eight documents have been associated with the "Technology" area ([17, 26, 40, 62, 69, 90, 92, 95]) and are studies focusing on using semantic ontologies to allow and promote the management of processes through the implementation of Internet of Things, robotics and sensors. Four articles [26, 40, 68, 92] concern the role of semantic technologies in IoT applications and services. The first two articles present the Semantic Sensor Network (SSN) ontology for describing sensors and their observations, the involved procedures, the studied features of interest, the used samples and the observed properties, as well as actuators. SSN was initially published by the W3C Semantic Sensor Network Incubator Group (SSNO). The current version of SSN is based on a revised and expanded version of the Stimulus Sensor Observation (SSO) pattern, namely the Sensor, Observation, Sample, and Actuator (SOSA) ontology. The ontology aims to represent sensors, their observations and all the concepts that revolve around this specific domain. SSN is very versatile and flexible, therefore applicable in a wide range of situations, like the management and control of wearable sensors for both employees and patients or the interconnection of devices. Cornejo et al. [17] present a complete ontology called OntoSLAM, developed to solve Simultaneous Localization and Mapping (SLAM) problems in different domains. Similar issues are also studied by Joo et al. [62] in their development of a scalable navigation framework for robots in various environments and scenarios, based on the Triplet Ontological Semantic Model (TOSM). SLAM computational problem of constructing or updating a map of an unknown environment while simultaneously keeping track of an agent's location within, it is a crucial problem when applying automated-guided robots inside the hospital environment, e.g., for automatic drug collection and delivery. The aspect of the implementation of technologies for assisting processes in the healthcare domain is explored by Santana et al. [95]. This study exhibits the methods and the results for designing the TEON, an ontology for the telehealth domain. TEON has been developed for obtaining a formal representation of the proper domain, such as second opinion, education, teleconsultation or telediagnosis, finding a way to let telehealth systems exchange data and integrate heterogeneous sources. The article offers a comparison

between the developed ontology and other studies centred on a semantic representation of the telehealth domain, as already existing medical and clinical vocabularies (e.g., SNOMED-CT) did not provide the terms to represent crucial concepts related to this specific domain, being unable to reach the high degree of formality that TEON did. The ontology was developed based on the upper domain ontology BioTopLite2 and Ontology for Biomedical Investigations (OBI) and built following the guidelines of a set of competence questions, regarding the individuation of the subareas of telehealth, the embedded services, the roles performed by the actors and the delivered processes. The main components of TEON are `Actors` (i.e. requestors of the service), `Teleconsultants`, `Manager`, `Services` (including the delivery of selected healthcare specialities), `Time` and `Space` classes and axioms.

The work by Prestes et al. [90], does not strictly concern healthcare, focusing on the introduction of the Core Ontology for Robotics and Automation (CORA), which is defined by the IEEE 1872-2015 standard [52]. The ISO/FDIS 8373 standard vocabulary has been adopted as one of the sources of domain knowledge for building the ontology [55]. The main aim of the ontology is to provide a semantic representation of the knowledge in the domain of robotics and automation. The result is a unified representation of a common set of definitions and relations that allow for the reasoning and communication of knowledge in this field. This ontology represents the fundamental concepts of the domain and serves as a base for more specific semantic representations. Its main concept is `Robot`, which is related to most of the remaining terminology through the sub-classes of `Device` and `Agent` (Fig. 3.5).

A set of six articles was identified as compliant with the topic of "Medical Vocabulary" [12,25,39,66,68,125]. All of the items are focused on ontologies that represent specific medical terminologies and the classification of terms related to the medical area. A formal semantic representation of the medical field is needed in every aspect concerning the progress towards the realization of a smart hospital environment. Hakimi et al. [39] aim to develop the Devices, Experimental Scaffolds and Biomaterials (DEB) ontology, a semantic representation of the domain of biomaterials. The ontology was created in order to research terms, enhance machine learning applications and provide a formal vocabulary of the domain. The reason why this semantic representation was developed was to have a tool that could cover all materials testes in a biological system to give a wider coverage of the terminology represented

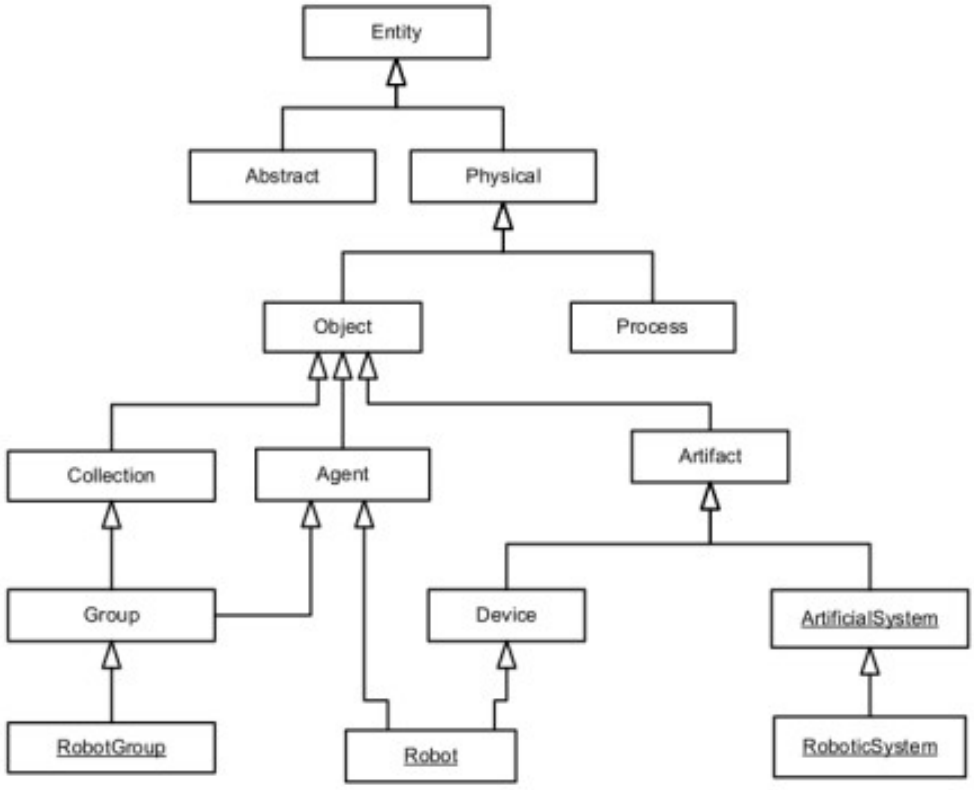


Figure 3.5: The CORA ontology [51].

and to complement other existing vocabularies. In DEB, a biomaterial is defined as "A non-drug raw material or substance suitable for inclusion in systems which augment or replace the function of bodily tissues or organs" and it is one of the superclasses of the ontology. The works of Bona et al. [12] and Liu et al. [68] aim to analyse the National Cancer Institute Thesaurus (NCIT), developed by the National Cancer Institute's Centre for Bioinformatics and Office of Cancer Communications with the main objectives of providing a base terminology for cancer, creating a vocabulary that is understandable by both humans and machines and promoting the introduction of new concepts and relationships derived from research, clinical trials and other information sources.

NCIT is a thesaurus that includes a broad coverage of the cancer domain, including cancer-related diseases, findings, and abnormalities. It is defined

as a controlled vocabulary organised as a list of terms and definitions. The ontology's domain includes vocabulary for clinical care, transitional and basic research, and administrative activities.

El-Sappagh et al. [25] studied a well-established standardised clinical vocabulary: the Systematized Nomenclature of Medicine Clinical Terms (SNOMED-CT).

It is a clinical healthcare terminology system used for electronic healthcare records. It includes concepts representing diagnosis, procedures, physical objects, body structures and many other information about health records (Fig. 3.6). The main component types are:

- `Concepts`, a numeric code with clinical meaning that is not human-comprehensible, but it is machine-readable;
- `Descriptions`, there are two types of descriptions, the FSN-Fully Specified Name which is a description of meaning, and the synonym;
- `Relationships`

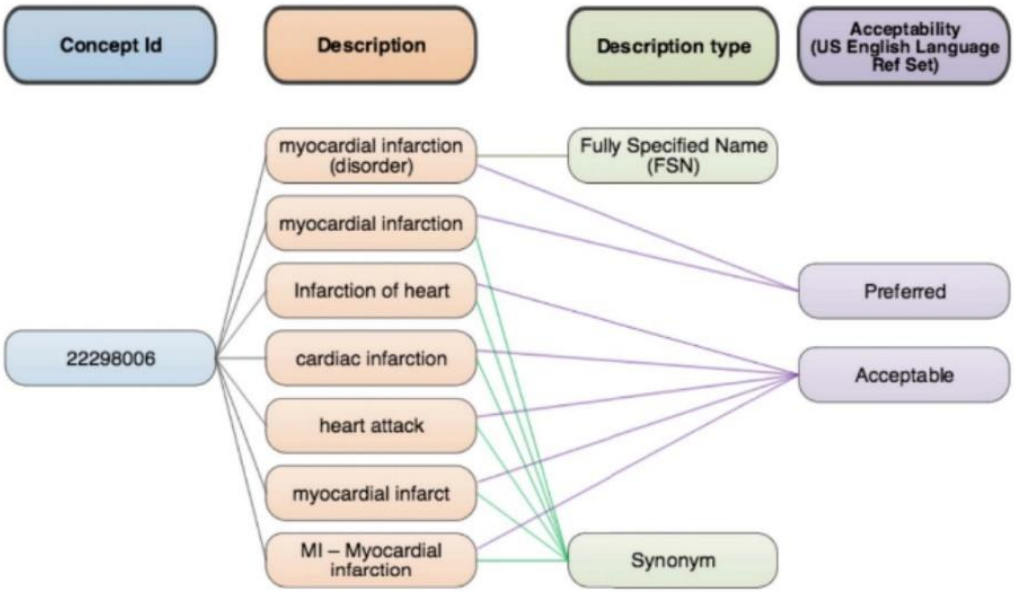


Figure 3.6: SNOMED-CT main types of components [98].

SNOMED-CT cannot be adequately represented through a semantic solution, due to inevitable issues that such translation would involve, which

are addressed by the article. For these reasons, the authors introduce the SNOMED-CT Ontology (SCTO). It is a standard ontology designed on the basis of the BFO and the Ontology for General Medical Science (OGMS). It is an upper-level ontology designed to represent the concepts of SNOMED-CT through a semantic representation. Concepts are implemented by adding further axioms and logical properties, providing a standard semantic representation that offers a wide coverage of the vocabulary items. SCTO can therefore be used in environments that support electronic data exchange, thanks to the logical semantics of the ontology format. The article by Kim et al. [66] is about developing the Dietary Lifestyle Ontology (DILON) with an extensible concept structure to support the interoperability of dietary lifestyle data from different cultural contexts. Dietary concepts and their relationships in DILON have proved to be useful for resolving the challenges introduced when treating an entire diet-related data element as a single concept. DILON can help extending the SNOMED-CT vocabulary as only 54% of dietary concepts of the former are mapped to the latter. Yu et al. [125] consider two specific ontologies concerning adverse events: the Ontology of Adverse Events (OAE) and the Ontology of Drug Adverse Events (ODAE).

OAE is a semantic representation that follows the OBO Foundry principles and that collects concepts suitable for monitoring adverse events of various types, aiming at improving and organizing adverse event information. An adverse event is defined as the negative event that follows a medical procedure and the ontology is designed to address this domain, without considering the processes that led to the event itself nor events derived from illnesses or diseases. It brings attention to the difference between adverse event and causal adverse event (Fig. 3.7): both occur after a medical intervention, but the second one, a subtype of the former, is used only if the event has certainly occurred as a result of the intervention itself. In addition, the ontology offers a representation of the factors that influence adverse event outcomes. ODAE describes and represents drugs, their chemical ingredients, adverse events and how these entities are related. It also follows the OBO Foundry principles, and it reuses terms from other existing ontologies, including OAE.

Six articles are about "Disease Vocabulary" [5, 27, 81, 94, 97, 102]. The former reviews the performance assessment of NCIT, SNOMED-CT, and Orphanet Rare Disease Ontology (ORDO) matching systems for FAIR (Findable, Accessible, Interoperable, Reusable) data. The aim of the study by

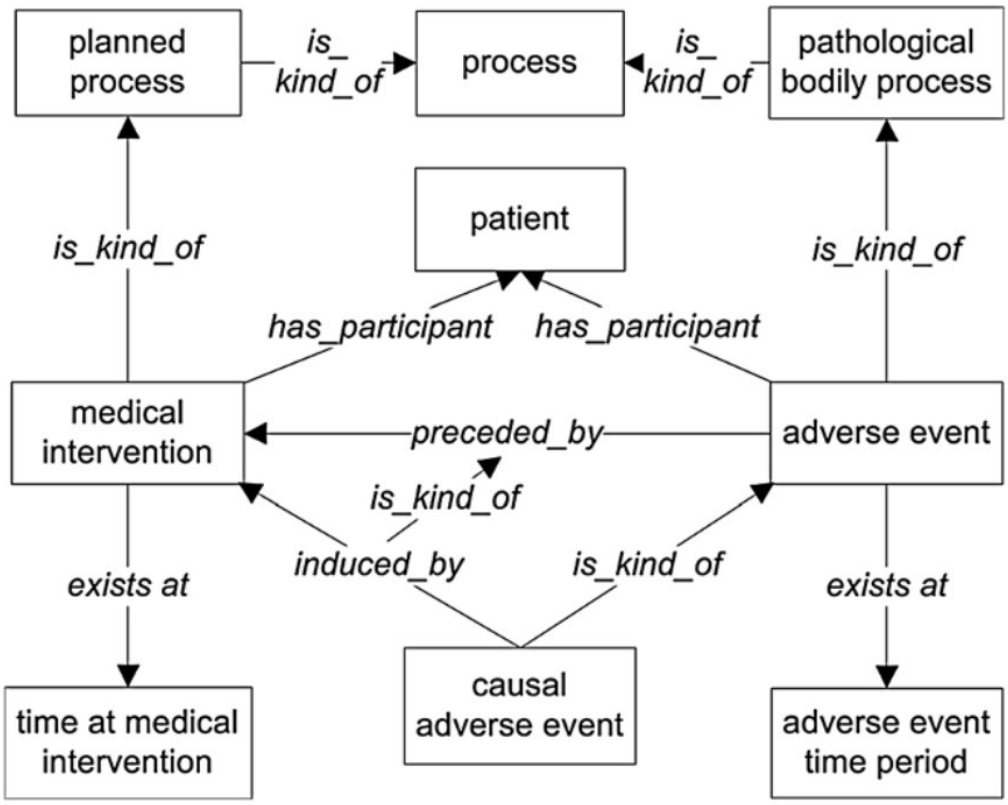


Figure 3.7: Basic design pattern of OAE adverse event and causal adverse event [43].

Esfahani et al. [27] is to provide an ontology for Multiple Sclerosis (MS) symptomatic treatment. According to the authors, a comprehensive ontological study addressing different concepts of MS symptomatic treatment is lacking. Therefore, the Symptomatic Treatment of Multiple Sclerosis Ontology (STMSO) has been developed with the objective of the study for building a knowledge base for developing Clinical Decision Support Systems (CDSS) in this domain. Silva et al. [97] study the application of ontologies and knowledge graphs in cancer research. It presents the aforementioned SNOMED-CT and NCIT, as well as the oncology subset (ICD-O) of the International Classification of Disease (ICD) and the Ontology for Biomedical Investigations, which aims to describe the terms related to biological and medical investigations. In regard to the NCIT, the authors involve an issue related to the discrepancy between most of the definitions included in the ontological form and the ones presented in the original thesaurus. Along

with this problem, the NCIT ontology also presents issues related to terms (e.g., problematic synonyms), and to the ontological representation in OWL, but nevertheless, it still provides a useful connection between the biological and clinical areas and a powerful tool to keep track of updates in these fields of interest. The article also analyses the ICD ontology. ICD is a classification system that organizes diseases and injuries into groups based on defined criteria. International Classification of Diseases 11th revision Clinical Modification (ICD11CM) describes in numerical or alpha-numerical codes the medical terms in which the diagnoses of disease or trauma, other health problems, causes of trauma and diagnostic and therapeutic procedures are expressed. The main classes of the ontology are `Diseases and Injuries`, and `Procedures`.

The article by Robinson et al. [94] concerns the Human Phenotype Ontology (HPO). The article focuses on the application of such ontology as a tool for analysing phenotypic abnormalities caused by hereditary diseases. The study of Narayanasamy et al. [81] reviews different ontologies for semantic-web applications in healthcare and virtual communications. Finally, the article by Babcock et al. [5] stresses the importance of the role of semantic representation as a powerful data-sharing tool when dealing with public health crises. The article gives a description of the Infectious Disease Ontology (IDO), which deals with the domain of infectious disease. IDO is based on the IDO-Core ontology, which takes a portion of its terminology from the OGMS and offers a general representation of the domain. It includes and defines several terms concerning the area of infections, such as infection, infectious disorder, infectious disease, and the process of establishing an infection. IDO also consists of the following IDO-Core extension ontologies (Fig. 3.8):

- VIDO represents an extension of IDO-Core and it is generally focused on the virus domain. Since it covers all the concepts in the domain of virus-induced diseases, it offers other IDO extensions, which also includes terminology that is already contained in the existing OBO Foundry.

- CIDO is the Coronavirus Infectious Disease Ontology. It offers a semantic tool that allows the representation of concepts related to this specific pathology, such as known and candidate anti-coronavirus drugs, genome data, host data, and vaccines. CIDO directly derives

from VIDO, adopting some of its terminologies and focusing on a specification of its domain. Although much more specific than its predecessor, CIDO concerns the coronavirus infectious diseases, therefore it includes all of the species of such viruses that can cause a large number of diseases.

- IDO-COVID-19 is the ontology, derived from CIDO, that specifically regards the domain of the COVID-19 disease and its cause SARS-CoV-2. It is still going through constant changes since the ongoing pandemic provides more and more items to be continuously adjourned.

The article also focuses on the problems which can originate from the application of such ontologies and their future improvement.

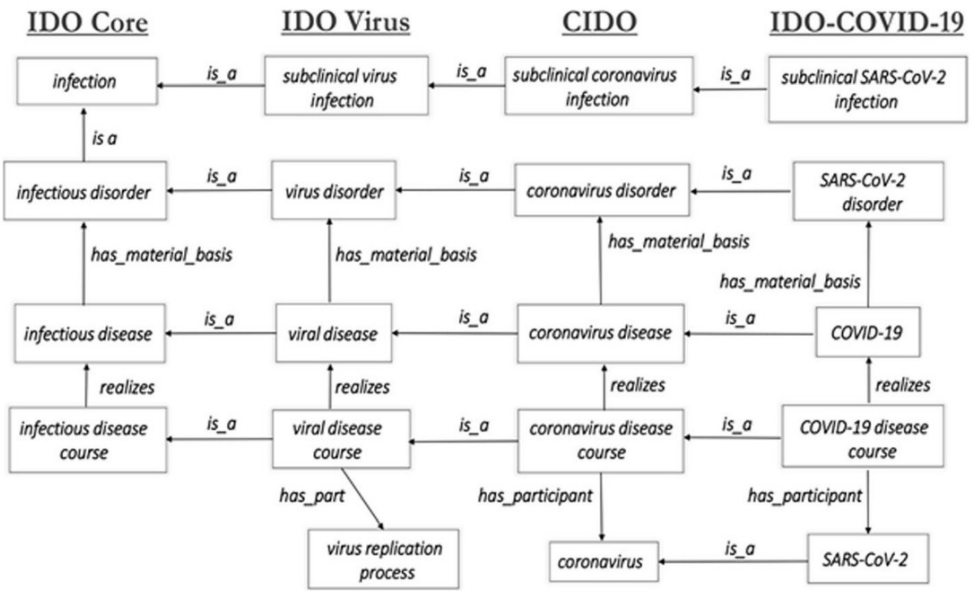


Figure 3.8: Links between VIDO, CIDO and IDO-COVID-19 ontologies [5].

Three selected articles are referenced as "Medical Data" [13, 59, 77]. The first article reports on the community effort to create the Data Management Plan (DMP) Common Standard Ontology (DCSO), with a particular focus on a detailed description of the components of the ontology. With the continuous growth of research data and the ultimate goal of sharing FAIR data, researchers face the challenge of systematically managing that data and its corresponding metadata. Data Management Plans make it easier for

researchers to respond to this challenge. DMP is a formal document that outlines how data are to be handled both during a research project and after the project is completed. The goal of a DMP is to consider the many aspects of data management, metadata generation, data preservation, and analysis before the project begins, which may lead to data being well-managed in the present, and prepared for preservation in the future. Therefore, the DCSO is taken into consideration within this review in relation to research being an integral part of medical activities and medical data management. The works by Ison [59] and McMurray [77] revolve around the description and study of ontologies that provide seamless exchange and collecting of medical data, with the purpose of enhancing interoperability between different healthcare structures and services. The former talks about the structure and scope of the EDAM (EMBRACE Data and Methods) ontology, whose main goal is to provide a semantic representation to identify and define the aspects of bioinformatics operations, which may also be understandable both by machines and humans. EDAM was developed for the EMBRACE (European Model for Bioinformatics Research and Community Education) project with the aim of offering a coherent, machine-understandable representation used within resource catalogues and to provide a common vocabulary for bioinformatics data and standards for data sharing. The main classes at the top of its hierarchy are: `Operation` represents how a piece of data is created; `Data` (which includes the additional sub-class `Identifier`), defines which data is consumed or produced by a tool; `Topic` includes the types of bioinformatics resource; `Format` for data formats. McMurray et al. [77] describe the actual lack of a system which is able to allow an effective information exchange between healthcare providers. In this regard, the Regional Healthcare System Interoperability and Information Exchange Measurement Ontology (HEIO) is proposed and described. HEIO has been designed with the specific purpose of enhancing the interoperability and information exchange among different healthcare providers and therefore obtaining a fully integrated healthcare system.

The only article in the collection that concerns the characteristics of the drugs domain is the work by Hanna et al. [41], which focuses on the process of building the Drug Ontology (DrOn), based on the standard drug terminology of the U.S. National Library of Medicine (NML) in RxNorm. The document goes through the process of the creation of the ontology itself, highlighting the building steps and the connection the developed ontology has with its

precursors. The following aspects are pointed out: extraction of data and information from RxNorm, transformation of such items into a Relational Database Management System (RDBMS) and the final translation in OWL. The article also provides descriptions of both the validation and the future plans of the developed ontology.

Three articles cover the domain of the "Human Role" in the health and clinical environment. This topic is essential because of the great relevance of human interconnections and the possible reachable complexity of inner organizations in any healthcare context. Hicks et al. [44] talk about the applications, the development, and the content of the Ontology of Medical Related Social Entities (OMRSE), which aims to semantically represent entities related to demographics, roles and characteristics of health workers. It is developed in OWL and defines gender roles, legal roles, healthcare providers and organization roles and patients. Being an OBO Foundry ontology, it reuses terminology from other existing representations, including BFO. Developers extended the domain over the years, adding specification classes to represent a wider variety of concepts, such as epidemic modelling. Maitra et al. [74] focus on the domain of interpersonal connection in medicine and the representation of data about presence in hospital structures through Presence Ontology (PREO). The document describes the survey, the domain literature review and the following steps which eventually lead to the creation of the ontology. The definition of classes and relational properties of found results is also performed, together with a final evaluation and description. Finally, the work by Gordon et al. [38] describes the development of a prototype knowledge graph, analysing the potential of semantic technologies to transform the idea of "geospatial open systems" into "open knowledge networks", which incorporate spatial and aspatial information across complex organizational networks. Ontology frameworks, such as VIVO, W3C Organization Ontology, Relation Ontology and *schema.org*, express the richness of relationships between organizations, projects and their collaborative work. Particularly, the Organization Ontology is a core ontology for organisational architectures and roles across a multitude of domains, and it can be used for representing all of the possible organisational interactions within the hospital. The areas represented by the ontology are the following: organisational structure, reporting structure (memberships, roles and relationships), location information, and organisational history. This representation does not offer specific details of the different

types of organizational structures, therefore, for this purpose, it is necessary to create extensions vocabularies. The Organization Ontology's classes are `Change Event`, `Formal Organization`, `Membership`, `Organizational Collaboration`, `Organizational Unit`, `Organization`, `Post`, `Role` and `Site`. All the above are then logically related through a multitude of properties.

Two articles are about "Buildings" [8, 22]. This topic is relevant in an accurate description of the healthcare environment, for example in terms of facility management as well as indoor localization and navigation. Donkers et al. [22] presents the Building Performance Ontology (BOP) which aims to enable the integration of topological building information with static and dynamic properties, to create a homogeneous data environment used by complex building performance assessments. Bassier et al. [8] offer an introduction to the Building Topology Ontology (BOT), with the analysis of its competence areas and applications in combination with other technologies. BOT originated from the need for the implementation of web-based applications to enhance the BIM methods. It defines the relationships between the components of a building and is used in the construction industry to promote the integration of linked data in the design, planning, construction, and maintenance of a building. The classes of the ontology follow:

- `Zone` is a part of the physical world or a virtual world that is inherently both located in this world and has a 3D spatial extent;
- `Site` is a part of the physical world or a virtual world that is inherently both located in this world and has a 3D spatial extent. It is intended to contain one or more buildings;
- `Building` is an independent unit of the built environment with a characteristic spatial structure, intended to serve at least one function or user activity [56];
- `Storey` is a part of the physical world or a virtual world that is inherently both located in this world and having a 3D spatial extent;
- `Space` is a part of the physical world or a virtual world whose 3D spatial extent is bounded actually or theoretically, and provides for certain functions within the zone it is contained in.

The class `Zone` is the main class of the BOT ontology, while `Site`, `Building`, `Storey` and `Space` are all sub-classes. By linking the classes and the object

properties of the BOT, it is possible to create a map, at a semantic level, of the building, which represents a significant aspect of hospital management, per se (Fig. 3.9). The current integration of BIM smart management systems in hospitals, also combining AI solutions with the infrastructures and facilities, can benefit from the integration of the BOT ontology. To this end, BOT can be the better ontology to describe hospital spaces.

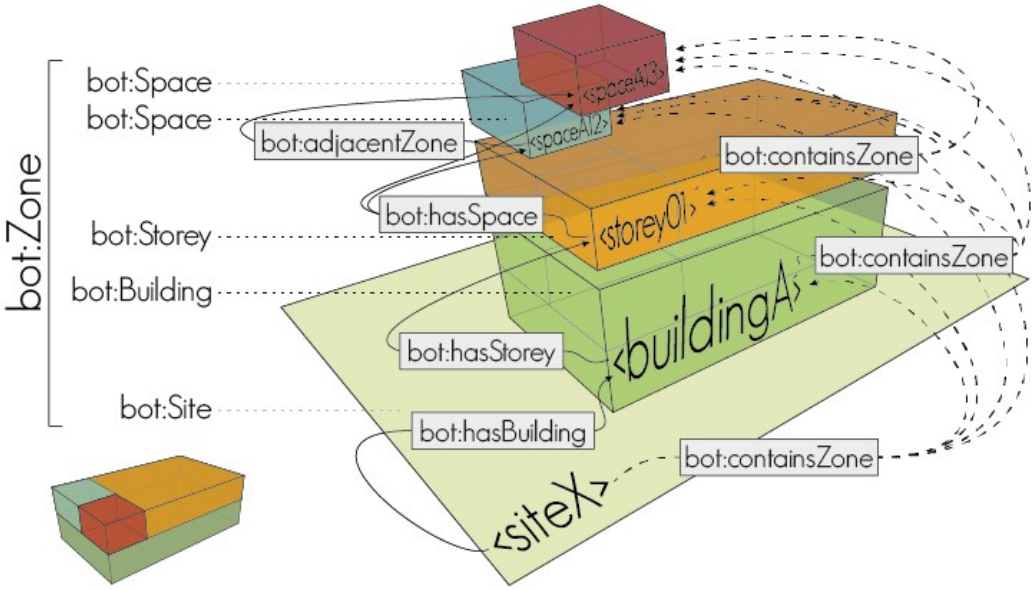


Figure 3.9: BOT Ontology - Examples of object properties linking classes [110].

The last three documents [36, 37, 76] concern a different area of interest each: "Services", "Medical Procedure", and "Emergency", respectively. Glockner et al. [37] explore the issue of lack of semantic representation of the logistics domain with the implementation of the Logistics Service Ontology Design Pattern (LoSe ODP). LoSe ODP describes the concepts linked to the logistics services area. It is an Ontology Design Pattern, which means that it is a small ontology that can be used as a base to design more specific ontologies. Ontology Design Patterns are used as a modelling approach to unravel issues related to ontology designs and reusability [35]. The competency questions that led the development process of the LoSe ODP revolved around some points of interest which needed to be covered with semantic representation: the actors involved in the providing of the service, the type of logistic service, the legal constraints related to the service, the required resources,

the information needed in the delivery of the service and the identification of the logistic service providers (LSP) together with the possible means of transportation. Many concepts represented by this ontology are taken from other ontologies regarding the field of the logistic supply chain. LoSe ODP reuses the notions of the differentiation between physical and informational resources, the importance of location according to the specific service provided, the objective and policies of the logistics service and the crucial role of time. The top level class of the ontology is `LogisticsService` which is logically related to the `Constraints` that has to sustain, the `Resources` that needs to consume in order to achieve its objective and the `Capability` of the logistics service that always involves a transformation.

The article by Gibaud et al. [36] covers the topic of Surgical Data Science (SDC) and the OntoSPM Collaborative Action. It states that information processing is strongly needed to perform surgical tasks and how the necessity for the creation of standardized Surgical Process Models (SPM) is relevant in such a scenario. Moreover, it also pinpoints that within IEEE there is a lack of appropriate regulations and standards for medical and surgical applications. The document explores and analyses OntoSPM, which has been developed in the context of the European initiative OntoSPM Collaborative Action [87], with the intent of developing ontology in the domain of surgical data science, both to create modelling scenarios from descriptions of real clinical cases and to have a tool that can be reusable in other contexts. OntoSPM focuses on SPM, actions and processes including roles played by the actors, affected objects (anatomy or pathology), instruments and materials and ways of manipulation. The article then proceeds to exhibit the development of two ontologies born as sub-ontologies of OntoSPM: the Ontology for Surgical Process Models in Laparoscopy (LapOntoSPM) and the Ontology for Data Integration in Surgery (ODIS). Current applications, as well as strategies to extend OntoSPM, including possible related issues, are also explored. The authors conclude by stating their strategy for ensuring medical acceptance, including the involvement of surgeons and the adoption of OntoSPM as a model for harmonization in surgical trials. One analysed topic which could benefit from the introduction of technologies and ontologies in the healthcare environment is the domain of emergency management during hazard crises. In such a scenario, having sufficient situational awareness information is critical. This requires capturing and combining data from sources like satellite photos, local sensors, and user-generated social media content. The

lack of an appropriate ontology that adequately conceptualizes this domain, gathers datasets, and integrates them, is a significant barrier to capturing, describing, and integrating such varied and diverse information. Mazimwe et al. [76] review numerous ontologies related to the disaster domain, such as Empathi, the Disaster Ontology, MOAC, Emergency Fire (EF), SMEM, SOKNOS, DOLCE. Among all the identified ontologies, Empathi, the Disaster Ontology, and DOLCE appear to achieve the better average score according to the implemented FAIR principles. Namely, Empathi has been designed with the aim of presenting state-of-the-art crisis vocabularies in order to conceptualise the core concepts of the management and planning of emergency response. The ontology has been developed importing concepts from already existing vocabularies concerning the hazard domain together with external ones that are not necessarily related to it. The resulting ontology contains super-classes that provide a generic coverage of the represented topic, such as `Event`, `Hazard`, `Type`, `Impact`, `Involved Actors` and `Services`. The broad set of concepts covered by Empathi makes it possible to extract structured information from sparse content, such as information coming from unstructured social media text. This semantic feature allows Empathi to enhance supervised and unsupervised learning in the crisis domain.

3.5 Conclusion

The chapter performs a brief analysis of the evident lack of an updated world standard for naming and coding medical devices and their fault codes associated with the maintenance work orders. This absence leads to clear issues when trying to collect data from different systems because mapping across different nomenclatures is nearly impossible due to the peculiar inner organisation of each nomenclature and CMMS software. This set of problems prevents the extraction of harmonised Real-World Data which is restraining, as a consequence, the development of Evidence-Based Maintenance which could otherwise provide guidance on improving the maintenance of medical devices while keeping medical equipment safe and reliable. In regards to the nomenclature of medical devices, something is moving, especially thanks to the efforts by WHO and the development of the International Classification and Nomenclature of Medical Devices. Instead, the standardisation of failure codes for maintenance is still in an embryonic phase as involved actors seem more inclined to propose new classifications from scratch rather than making

existing methodologies interoperable. Besides, they are also still attached to legacy approaches which, for instance, do not take software failures into account in spite of the ongoing spread of Health Information Technologies (see Section 2.2).

The scoping review identified 32 studies on the use of semantic ontologies to map different aspects of the healthcare environment. Studies have been classified into ten areas of interest: eight documents are associated with technology area, six studies are about disease vocabulary, six articles focus on medical vocabulary, three works relate to medical data, three papers cover the domain of human role, two are about building management, and the remaining four ones are each about drugs, services, medical procedures, and emergency.

A set of 34 ontologies extracted from the identified articles has been also analysed and discussed. A subset of the extracted ontologies laid the foundations for the development of the ODIN Ontology and the OdinEMDN Ontology within the European Project ODIN. Although the review's primary objective is not to analyse the FAIRness of the ontologies, it is interesting to observe that only seven out of the 32 identified papers make reference to the FAIR principles [13, 38, 39, 76, 97, 102, 125]. Despite all articles agreeing on the fact that ontologies lead to reproducible research and may improve the adoption of FAIR principles by supporting data integration, analysis, facilitating data interpretation, interoperability, and data mining, it emerges that appropriate metrics to evaluate the FAIRness are still developing. As Wilkinson et al. [118] state, the FAIR principles are aspirational, in that they do not strictly define how to achieve a state of FAIRness, but rather they describe a continuum of features, attributes, and behaviors that will move a digital resource closer to that goal. As a conclusion, the usefulness of semantic ontologies lies in the possibility of providing a suitable level of abstraction for sharing and reusing concepts in a standardised way so that the data gathered from heterogeneous sources receive a common nomenclature, empowering communications among actors and easing Healthcare Technology Management.

Chapter 4

Health Information Technology Adverse Events Identification and Classification with Natural Language Processing

In this chapter, the designing and developing of the proposed framework for facilitating the extraction of Real-World Evidence (RWE) using Natural Language Processing (NLP) and Explainable Artificial Intelligence (xAI) is illustrated. The framework is employed for the identification and categorization of adverse events related to Health Information Technologies (HITs) sourced from diverse Real-World Data (RWD) origins. In pursuit of this objective, adverse event records obtained from the US MAUDE database have been categorized by domain experts as either HIT-related or non-HIT-related, subsequently serving as the basis for fine-tuning a pre-existing model designed for binary text classification. The model's performance has been rigorously assessed through a 10-fold validation process and subsequently evaluated against a subset of records. Additionally, XAI techniques have been applied to elucidate the most prominent features contributing to each classification, thereby enhancing user comprehension of HIT-related adverse events.

4.1 Materials and Methods

4.1.1 Proposed framework

The proposed framework is illustrated in Fig. 4.1. Input data consists of medical device adverse event reports extracted from the US MAUDE database.

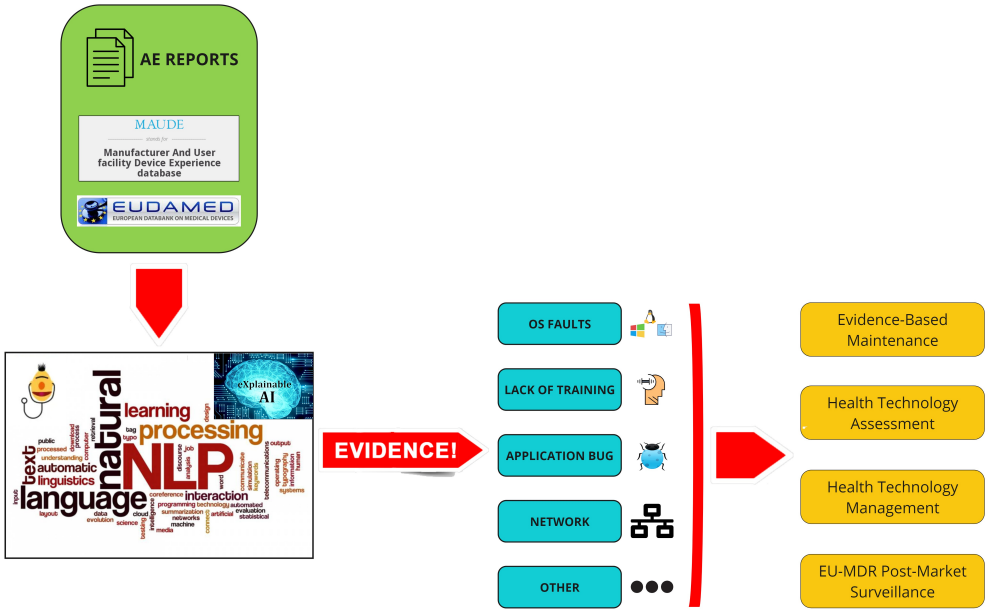


Figure 4.1: Proposed framework.

The developed model is based on the emilyalsentzer/Bio_ClinicalBERT model from HuggingFace, initialized from BioBERT (BioBERT-Base v1.0 + PubMed 200K + PMC 270K) and trained on all notes from MIMIC III, a database containing Electronic Health Records (EHRs) from Intensive-Care Unit patients at the Beth Israel Hospital in Boston, MA [3]. The model uses 12 layers of transformers block with a hidden size of 768 and a number of self-attention heads as 12. The pre-trained Bio_ClinicalBERT model has been fine-tuned on 3,705 manually-labelled adverse events reports extracted from the MAUDE database, to help the model learn domain-specific knowledge and terminology, leading to more accurate predictions. The implemented model performs binary text classification between HIT and non-HIT adverse event reports. XAI is also applied to the model to understand the weights

of each feature related to the output classes. Weighted keywords extracted from both output classes may also be used to help users to label new records, providing a prediction classification score. The developed framework represents a novelty in the Clinical Engineering field, as the RWE extracted can be applied for EBM, HTM, HTA, and PMS scopes as mentioned in Section 1.

4.1.2 Dataset statistical analysis

1,857 reports extracted from the FDA MAUDE were labelled by experts as HIT-related between January, 1st 2008 and June, 30th 2010. The remaining 513,183 reports which belonged to the same time span were otherwise classified as non-HIT. After discarding the records with no narrative data associated, the resulting dataset contained 492,030 records. 1,848 reports have been randomly sampled from the non-HIT population in order to build a balanced training dataset for the BERT classifier (class weights are respectively 1.0132 and 0.9871 for HIT and non-HIT adverse event reports). Kolmogorov-Smirnov test has been performed to compute the distances between the empirical (sample) and the theoretical (original) distributions and check whether the two follow the same distribution in relation to the manufacturer and the medical speciality (Fig. 4.2).

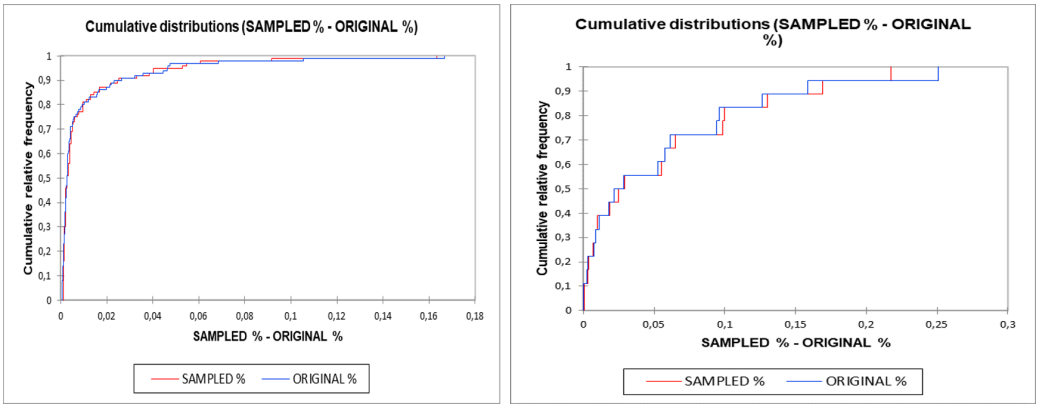
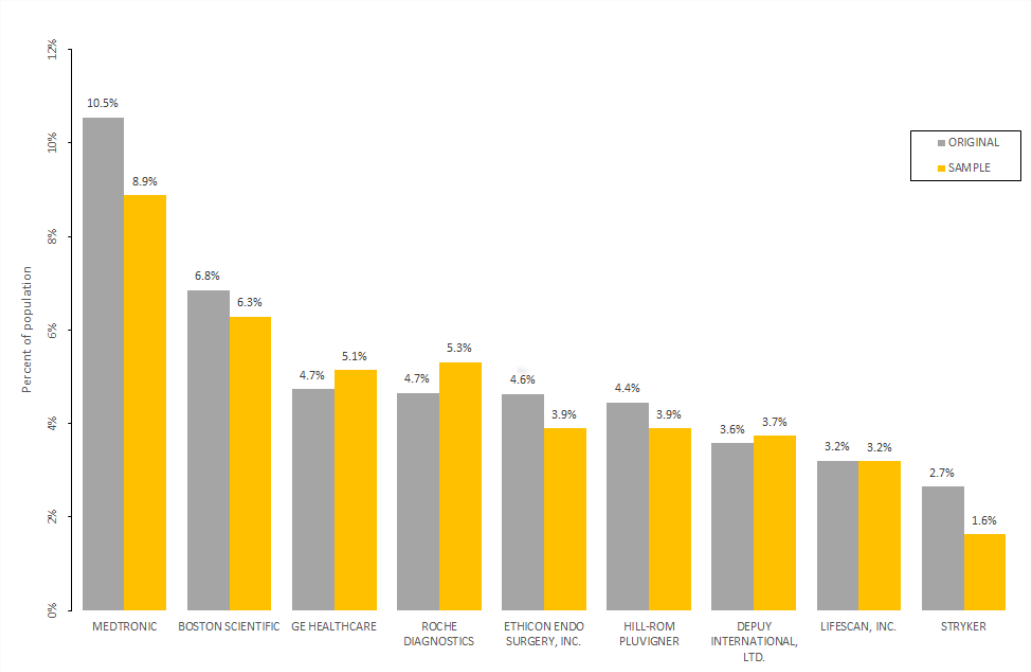
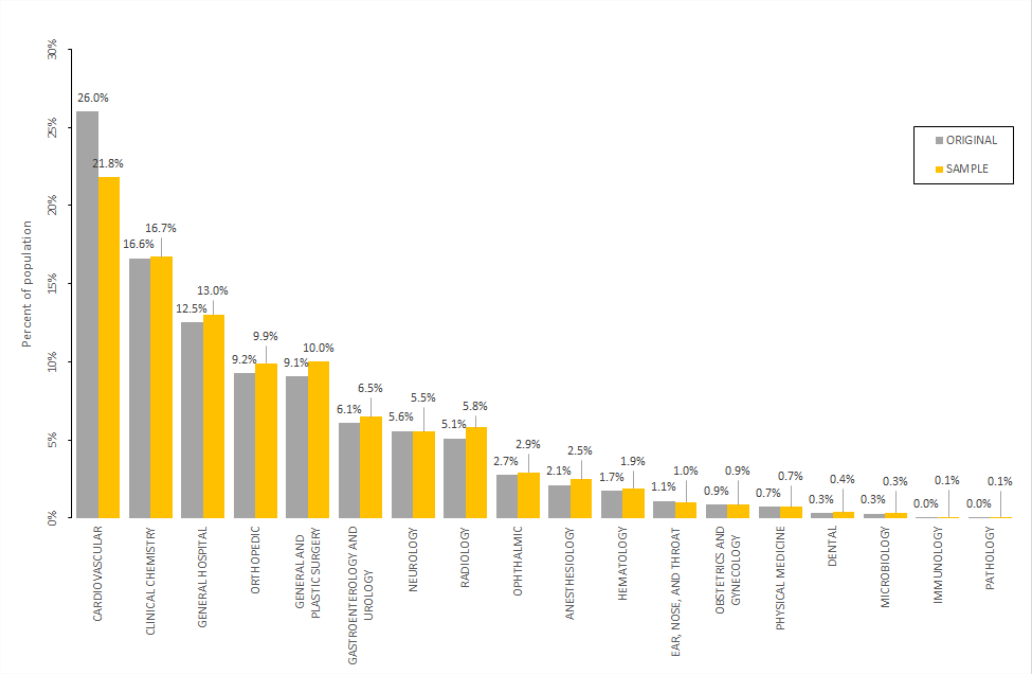


Figure 4.2: Results of Kolmogorov-Smirnov test applied to non-HIT data grouped by manufacturer (left) and medical speciality (right): p-values are 0.281 and 0.846 respectively.

Manufacturer and medical speciality categorical variables were statistically described as frequency count and percentage (Fig. 4.3).



(a) Bar chart of the percentage of original and sampled datasets for the top-10 manufacturer classes.



(b) Bar chart of the percentage of original and sampled datasets for the identified medical specialities.

Figure 4.3: Bar chart of the percentage of original and sampled datasets.

Texts longer than 512 words have been truncated without losing any meaningful information as they represent 99.14% of the whole dataset (Fig. 4.4).

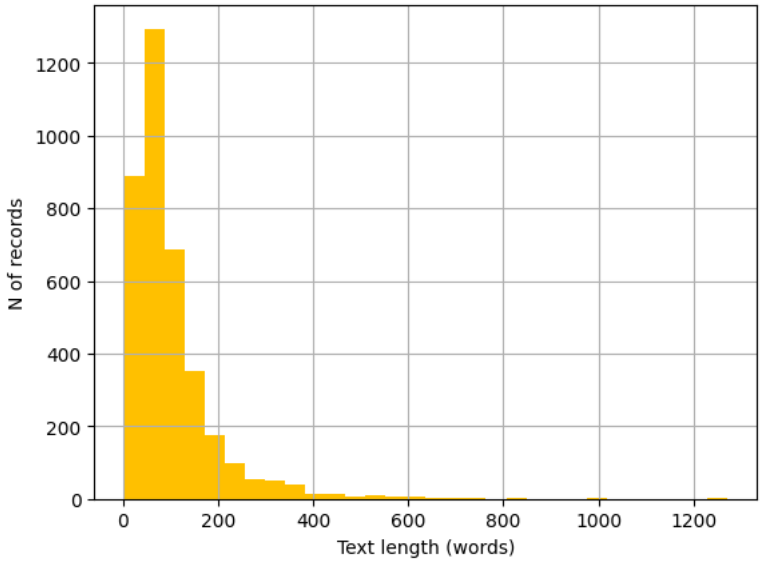


Figure 4.4: Number of words for analysed records. The majority of records (99.14%) present a text length which is shorter than 512 words.

The combined reports have been used for training, evaluation, and testing on an NVIDIA GeForce RTX 3090. The model has been tested on 741 never seen records (20% of the whole dataset). The remaining samples (2,964 rows) have been split respectively into 80% for training and 20% for validation.

4.1.3 Explainable Artificial Intelligence

Two surrogate XAI models have been used to understand which are the main features that affect the output of the model, in order to unravel the decision-making process: LIME and SHAP.

LIME stands for Local Interpretable Model Agnostic Explanation. The "local" aspect means that it is used to explain individual predictions of a machine learning model. Each text record within the test set is explained in terms of keywords, each one weighted in terms of relevance to the contribution to the final binary classification [93].

SHAP (SHapley Additive exPlanations) is a method based on cooperative game theory and used to increase the transparency and interpretability of machine learning models. SHAP value shows how each feature affects each final prediction, the significance of each feature compared to others, and the model's reliance on the interaction between features. To evaluate an existing model f when only a subset S of features are part of the model, the other features are integrated by using a conditional expected value formulation:

$$E[f(X)|X_s = x_s]$$

When explaining a prediction $f(X)$, the SHAP value for a specific feature i is the difference between the expected model output and the partial dependence plot at the feature's value x_i [70].

4.2 Results

Various experiments have been initially conducted in order to tweak the model parameters. Dropout hyper-parameters have been constantly set to 0.5 for the attention layer and 0.1 for the hidden layer [24]. Three different activation functions - the Sigmoid Linear Unit (SiLU), the Rectifier Linear Unit (ReLU), and the Gaussian Error Linear Unit (GELU), three learning rates for the optimization algorithm ($5e^{-5}$, $3e^{-5}$, and $2e^{-5}$), and three batch sizes (8, 16, and 32) have been tested as suggested by Devlin et al. [20]. Tests have been also conducted on the number of frozen layers to achieve the best performance: 0, 4, 8, or 12 encoder layers have been frozen. A test has been also conducted by only freezing the embedding layers. The best performances have been achieved with 8 frozen layers, the GELU activation function, a batch size of 16, and the AdamW optimizer with a learning rate of $2e^{-5}$. The model has been trained for a total of 30 epochs. Fig. 4.5 plots the comparison between training and validation loss, showing how the model begins to overfit after seven epochs.

The observed trend is coherent with the general approach of fine-tuning BERT-based models for just a limited number of epochs [20]. Therefore, the model has been trained only for seven epochs to avoid overfitting, obtaining 0.9680 accuracy, 0.9603 precision, 0.9764 recall, and 0.9683 F1 score. Finally, a new model has been trained on the same dataset with the same tuned hyper-parameters and k-fold cross-validation (10 folds). Results are shown in Table 4.1.

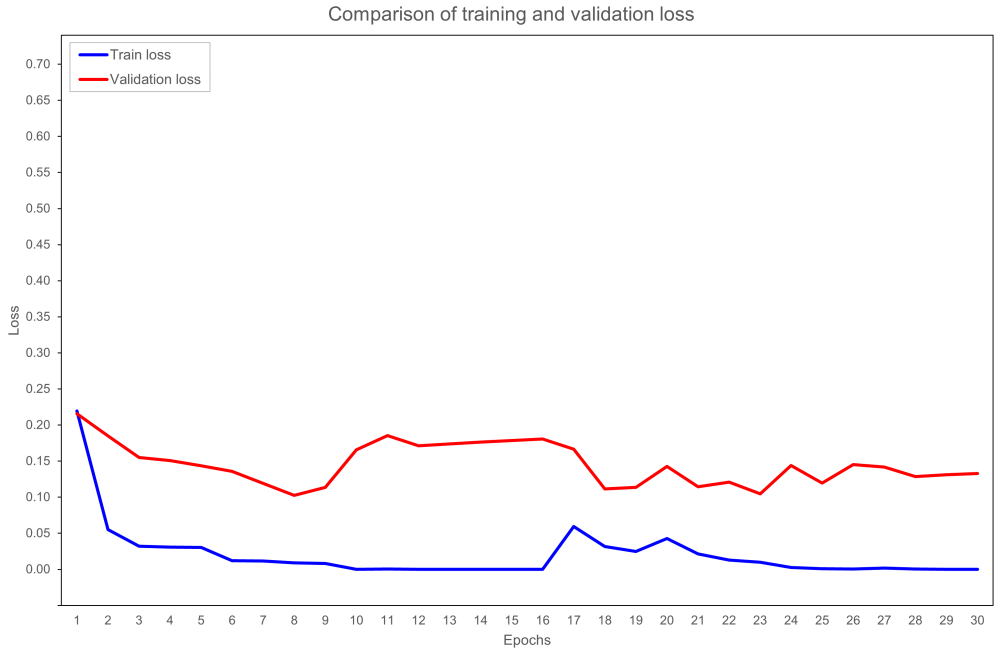


Figure 4.5: Comparison of training and validation loss during 30 epochs of training.

Table 4.1: Results of 10-fold validation on 2,964 records. Highlighted fold 4 shows the best overall metrics.

Fold	Train loss	Validation loss	Accuracy	Precision	Recall	F1 score
1	0.0385	0.1277	0.9663	0.9860	0.9463	0.9658
2	0.0145	0.0169	0.9966	1.0000	0.9933	0.9966
3	0.0026	0.0102	1.0000	1.0000	1.0000	1.0000
4	**0.0001**	**0.0012**	**1.0000**	**1.0000**	**1.0000**	**1.0000**
5	0.0002	0.0088	1.0000	1.0000	1.0000	1.0000
6	0.0001	0.0028	1.0000	1.0000	1.0000	1.0000
7	0.0002	0.0026	1.0000	1.0000	1.0000	1.0000
8	0.0037	0.0047	0.9966	0.9933	1.0000	0.9966
9	0.0002	0.0143	1.0000	1.0000	1.0000	1.0000
10	0.0001	0.0014	1.0000	1.0000	1.0000	1.0000

4.2.1 Testing and performance evaluation

The best-performing model (fold number 4) has then been tested on the testing dataset, with the following performances: 0.9946 accuracy, 0.9893 precision, 1.0000 recall, 0.9946 F1-score, and 98.93% Matthews Correlation Coefficient. Fig. 4.6 compares the training and validation loss in relation to the training epochs to ensure that there is no overfitting during training. The Receiver Operating Characteristic (ROC) curve and confusion matrix for testing results are shown in Fig. 4.7.

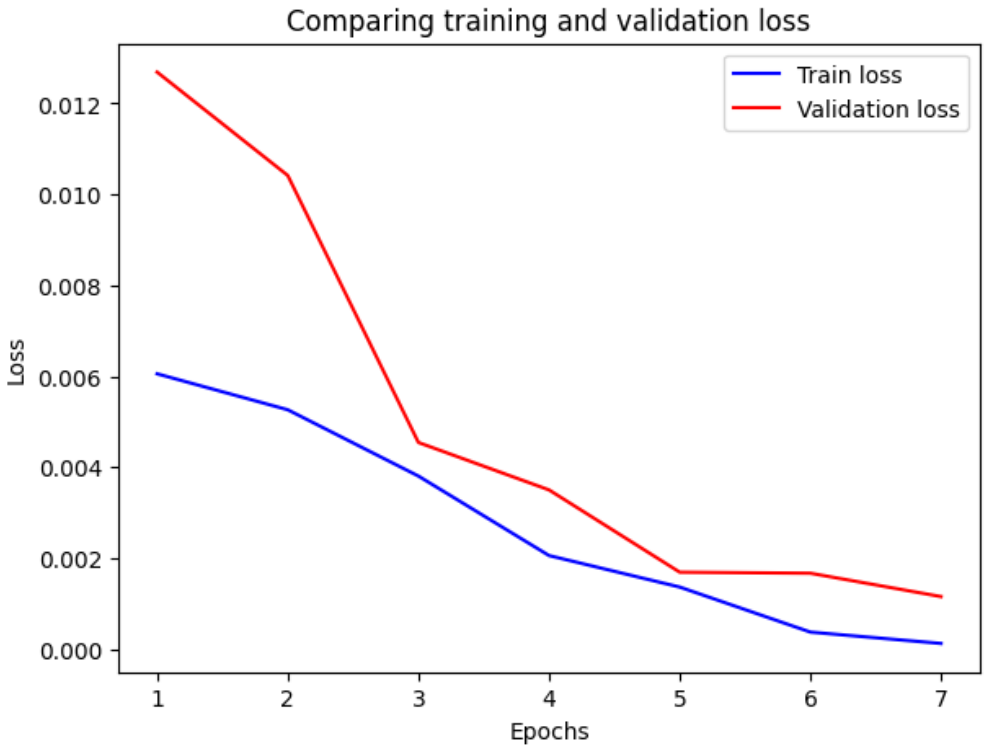


Figure 4.6: Comparison of training and validation loss for fold 4. Both losses decrease during the epochs so that it can be asserted that there is no overfitting.

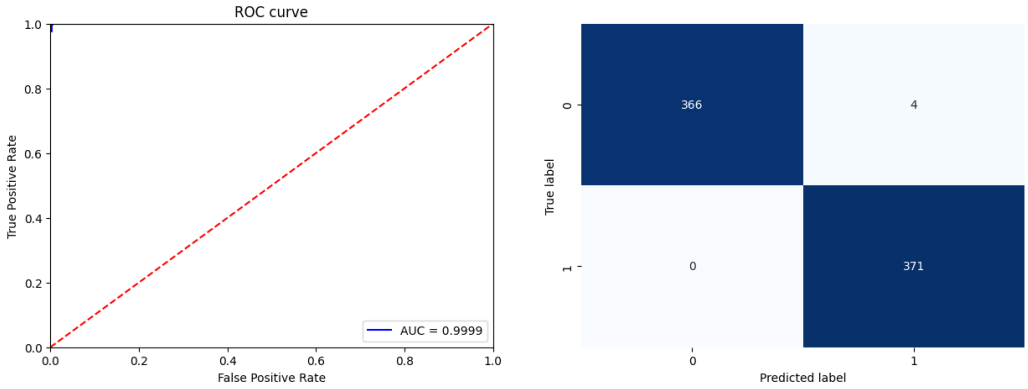


Figure 4.7: ROC curve and confusion matrix for fold 4 tested on 741 records (20% of the whole dataset).

The developed model has an overall classification run-time of 9.73 s ± 21.5 ms for 1,000 reports. The classification run-time of one report is 9.48 ms ± 5.6 μs. The Python code, the training and validation datasets, as well as the final fine-tuned model, are available on a public GitHub repository [1]. Results show better metrics than other existing HIT adverse events reports text classifiers based on non-BERT NLP models (Table 4.2).

Table 4.2: Comparison of performances of the proposed NLP model (fine-tuned ClinicalBERT) and other non-BERT models. LR - Logistic Regression. SVM - Support Vector Machine. CNN - Convolutional Neural Network. HRNN - Hierarchical Recurrent Neural Network.

Model	Accuracy	Precision	Recall	F1 score
ClinicalBERT	**0.9946**	**0.9893**	**1.0000**	**0.9946**
LR [15]	-	0.9670	0.9420	0.9540
LR [33]	-	0.6940	0.8040	0.7450
SVM+LR+CNN [116]	0.9012	0.8796	0.8606	0.8700
LR+CNN+HRNN [63]	0.9030	-	-	0.8760

[1] https://github.com/alessioluschi/HITBert

4.2.2 Explainable AI applied to the model

Fig. 4.8 shows the bar plot of the top 20 features obtained with SHAP applied on the best-performing model (fold 4) on the test set. Words such as "handheld", "computer", "screen", and "software" have a high positive contribution to the prediction of the HIT class, while "device", "product", and "reported" have a negative contribution to the prediction, reflecting a positive weight for the non-HIT class.

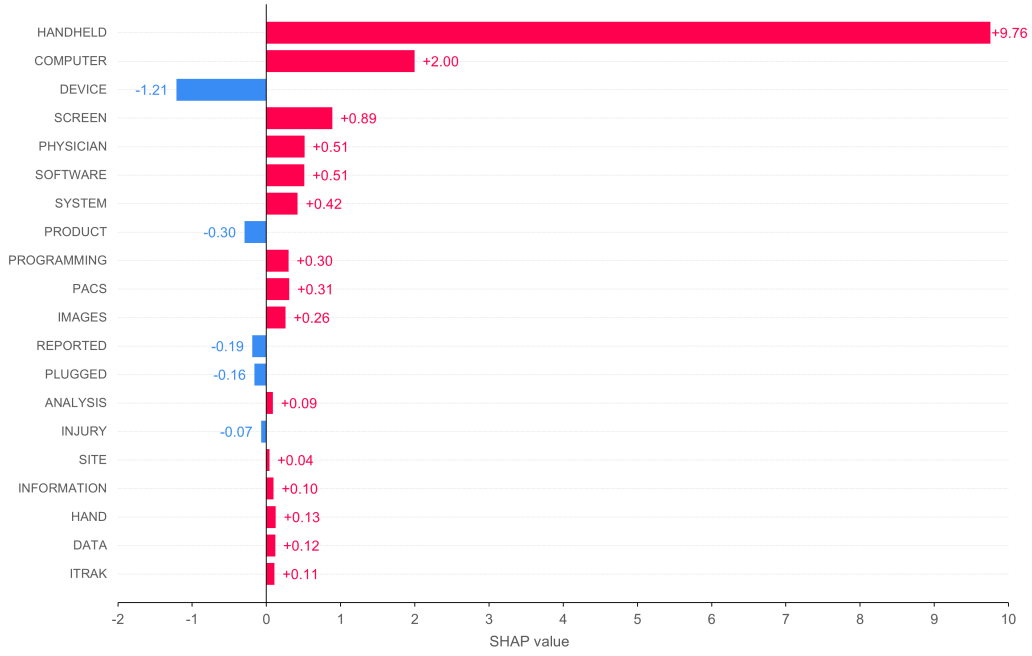


Figure 4.8: Bar plot of the top 20 features analysed with SHAP.

Fig. 4.9 shows how LIME explains the output classification for a given text with the top 10 features: words such as "track", "tracker", or "system" have a high weight related to the HIT output class, so they are mainly responsible for the final classification of the model (which in this case is concordant with the label).

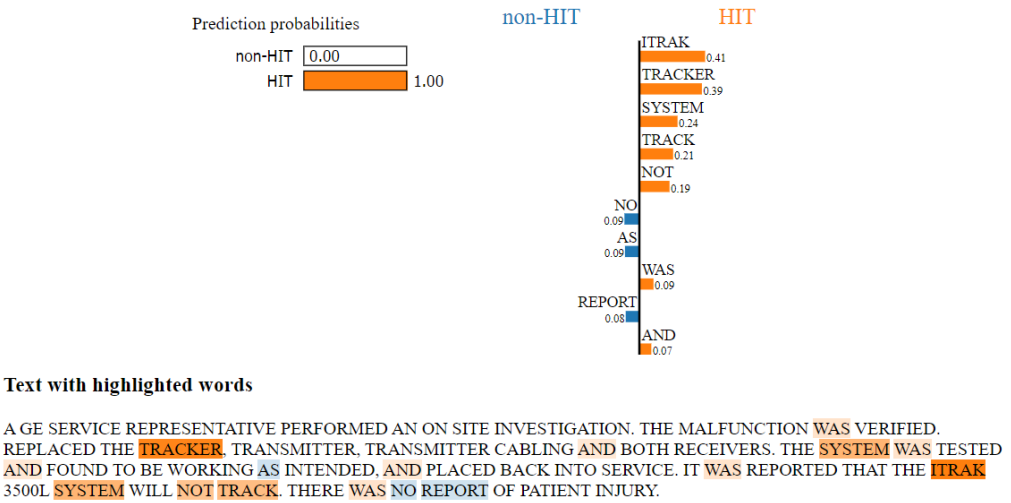


Figure 4.9: LIME applied to record with MDR 978358 for top 10 features. Words like "track", "tracker", and "system" have a strong significance for the HIT output class, and they are those responsible for the final model classification.

The overall weight for each keyword is calculated according to the formula:

$$F_i = \frac{\sum_{n=1}^{k} m_n}{k}$$

For each keyword, all the magnitudes identified within the texts m_n are summed together and then divided for the total count of the analysed keyword k. The calculated value F_i is the overall weight associated with the given keyword, representing the overall importance of that feature within the whole testing set in relation to the final classification. The process is applied both to HIT and non-HIT keywords (Fig. 4.10 and Fig. 4.11).

Words like "handheld", "itrak", "ultrasound", or "centricity" are associated with high average magnitudes in relation to the HIT class. Words like "fowler", "bilirubin", "bladder", or "reprocessed" have instead higher average magnitudes for the non-HIT class.

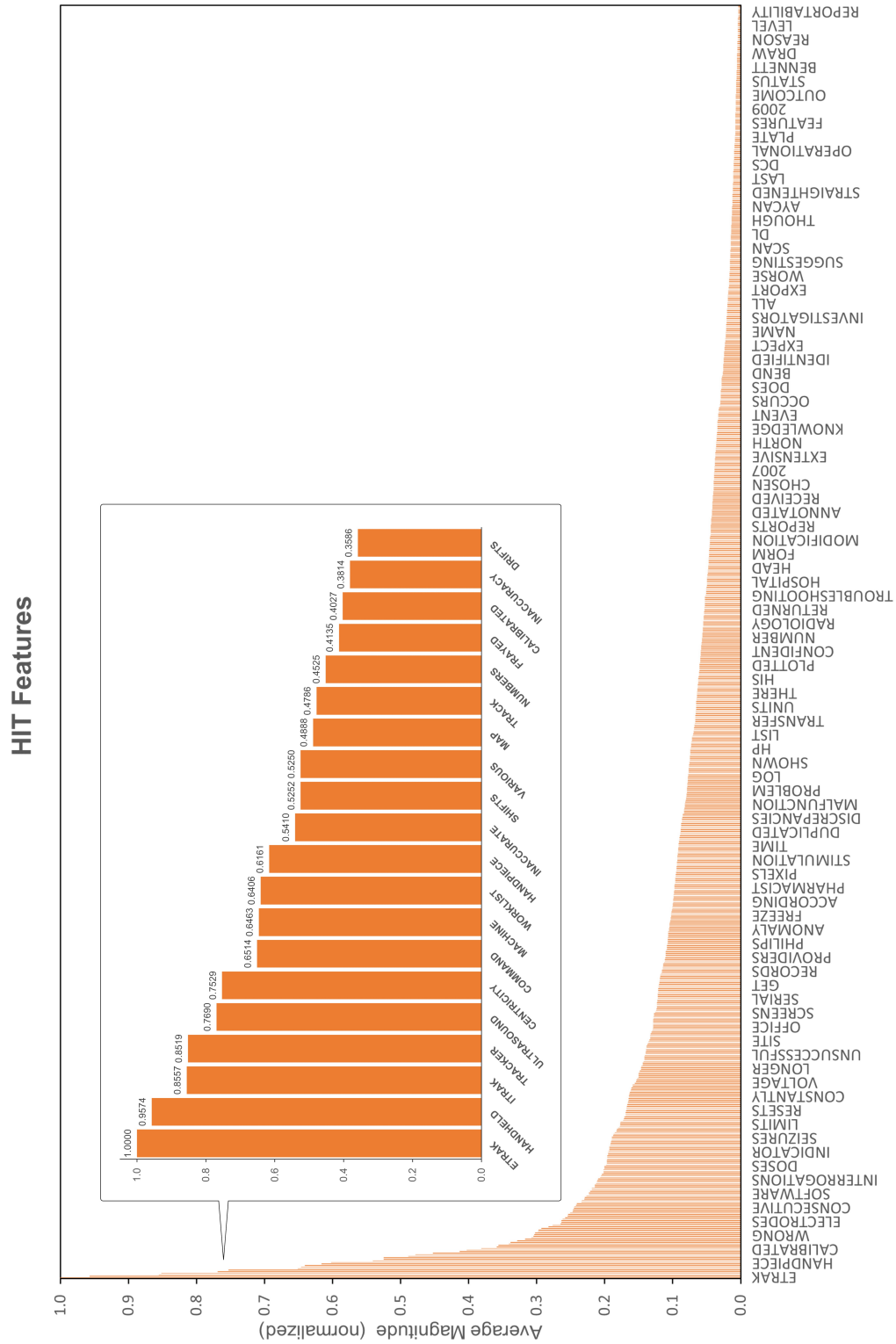


Figure 4.10: Bar plot for the keywords (features) related to HIT classification in relation to the average normalized weight, extracted with LIME. The top 10 features are highlighted in the callout.

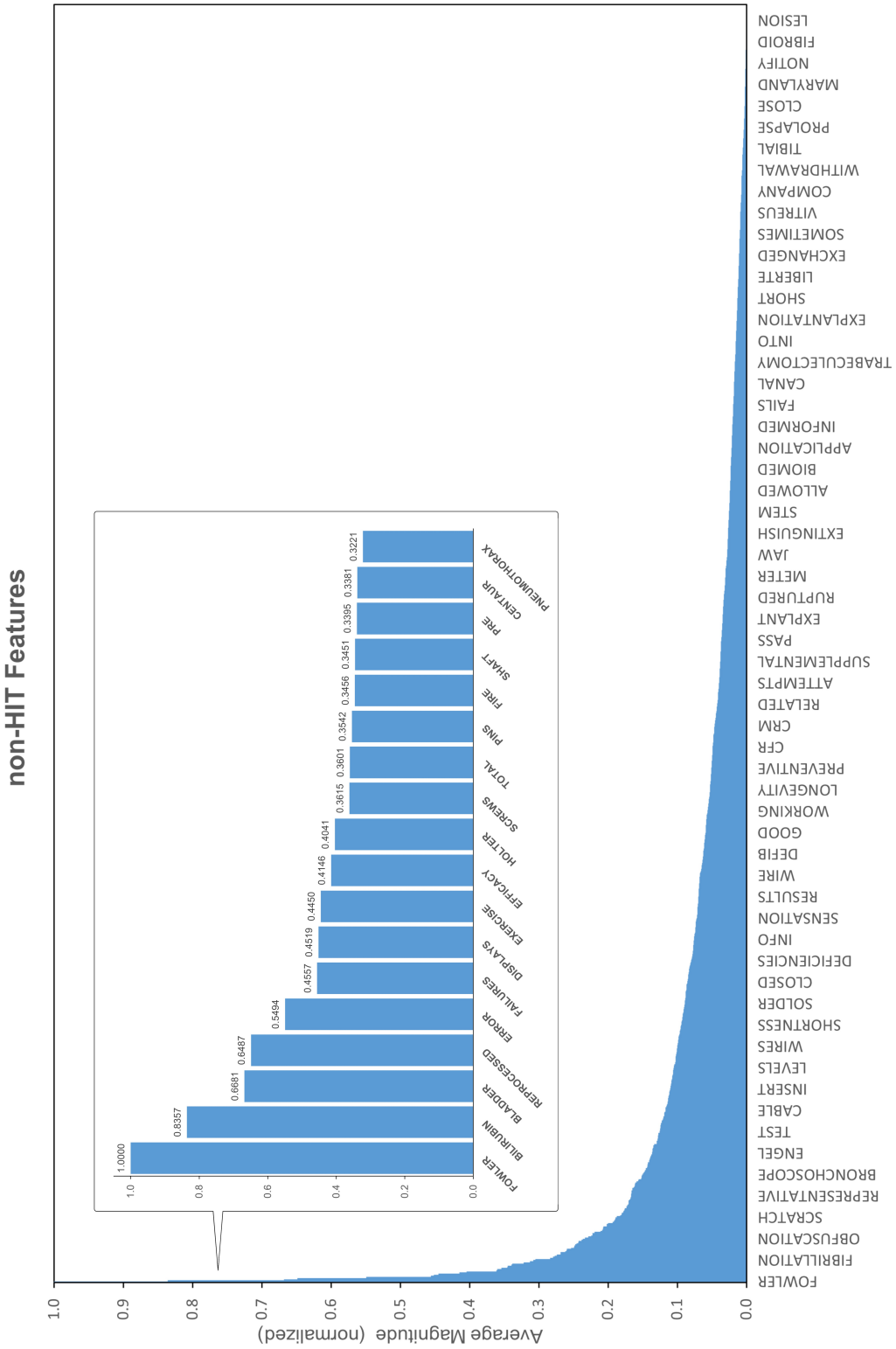


Figure 4.11: Bar plot for the keywords (features) related to non-HIT classification in relation to the average normalized weight, extracted with LIME. The top 10 features are highlighted in the callout.

4.3 Discussion

4.3.1 Implementation and best practice

The definition of the hyperparameters of the model followed the common best practice for fine-tuning BERT models as described by Devlin et al. [20]. All the tested values for the activation function, the batch size, the optimizer, and the learning rate are those suggested by the authors of the language representation model itself. Another common best practice is to use only a few epochs for fine-tuning BERT models for domain-specific tasks, as a pre-trained model usually requires a much smaller number of epochs than models trained from scratch. In fact, the authors of BERT recommend between two and four epochs [20]. Further training often translates to overfitting the data and forgetting the pre-trained weights. This phenomenon, termed catastrophic forgetting, may occur when new data contrasts with what the Large Language Model has previously learned, resulting in overwriting or dampening the old information. A high variability in accuracy between runs with the same settings has also been observed. This instability has been known since the release of BERT. While catastrophic forgetting and the small sizes of the dataset were first suspected as the causes of this instability, more recent work [79] suggests that optimisation difficulties leading to vanishing gradients are the actual reasons. This represented a huge issue during the development of the model, even though defining fixed seeds and implementing layer-freezing during the fine-tuning phase seemed to partially mitigate the problem.

4.3.2 Explainable AI

XAI can help understand the process behind the final classification performed by the model. The two implemented strategies (SHAP and LIME) bring similar results for the most relevant features which led the model to the HIT class, while the results for the non-HIT class are different. The discrepancy may be mainly due to the different types of variables analysed in the charts. A single feature may have a not-high cumulative SHAP value (reflected in a not-high general weight) but with a high average magnitude due to its low frequency in the whole analysed corpora. It is also relevant to analyse what brought the model to perform an incorrect classification. We discuss the extracted texts analysed with LIME related to the four false

positives classification (Fig. 4.7). In the first case, it clearly emerges that the word "handpiece" is the main responsible for the misclassification, because it has been commonly associated with HIT adverse events by the model during the training phase (Fig. 4.12).

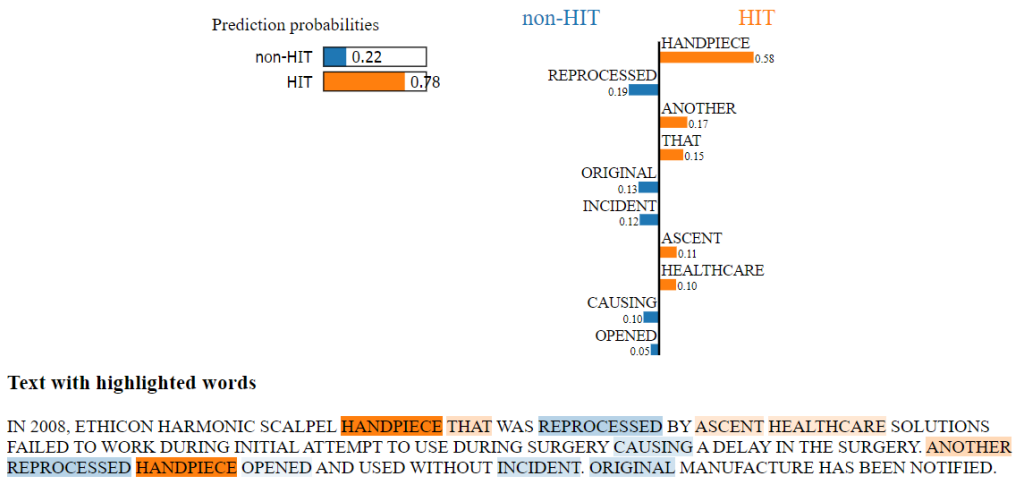
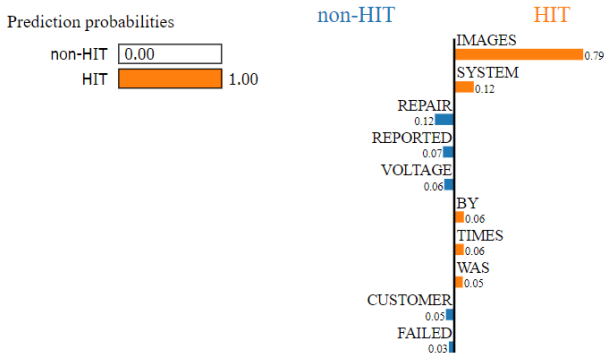
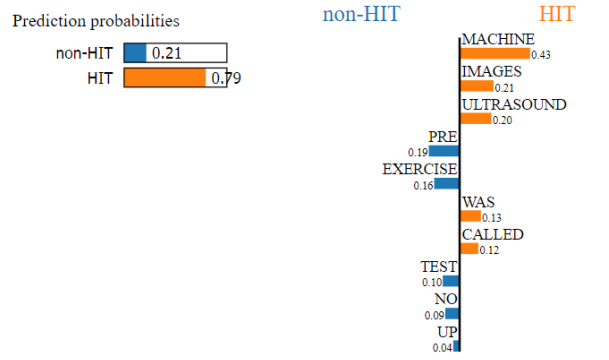


Figure 4.12: LIME applied to one adverse event report classified as HIT but labelled as non-HIT by experts. The misclassification is mainly due to the word "handpiece".

The second and the third reports (Fig. 4.13a and Fig. 4.13b) are totally different cases. In fact, for these reports, the initial label assigned by the experts is actually wrong, and the texts are correctly associated with HIT adverse events: a mix-up of images due to a communication error in the first example, and a system freeze issue during pre-exercise image acquisition in the second one.



(a) LIME applied to one adverse event report classified as HIT but labelled as non-HIT by experts. The words "images" and "system" strongly contribute to classifying the record as HIT adverse event. By reading the text, it can be assumed that the report is about a communication error with the subsequent mix-up of images, which can be actually considered a HIT-related adverse event.



(b) LIME applied to one adverse event report classified as HIT but labelled as non-HIT by experts. Words such as "machine", "images", and "ultrasound" lead the model toward a HIT prediction, which is actually true, being the initial label wrongly assigned as it emerges by reading the text.

Figure 4.13: LIME applied to reports with MDR 1182990 and 1028497.

In the last example (Fig. 4.14), it can be observed that the features with higher weights such as “numbers”, “34”, and “various” lead the model toward the wrong classification, as the report is related to a medication cassette reservoir. Additional fine-tuning of the model should be performed in order to make it able to contextualise the words “numbers” and “various” more specifically, resulting in lesser associated classification weights as they are more common speech features.

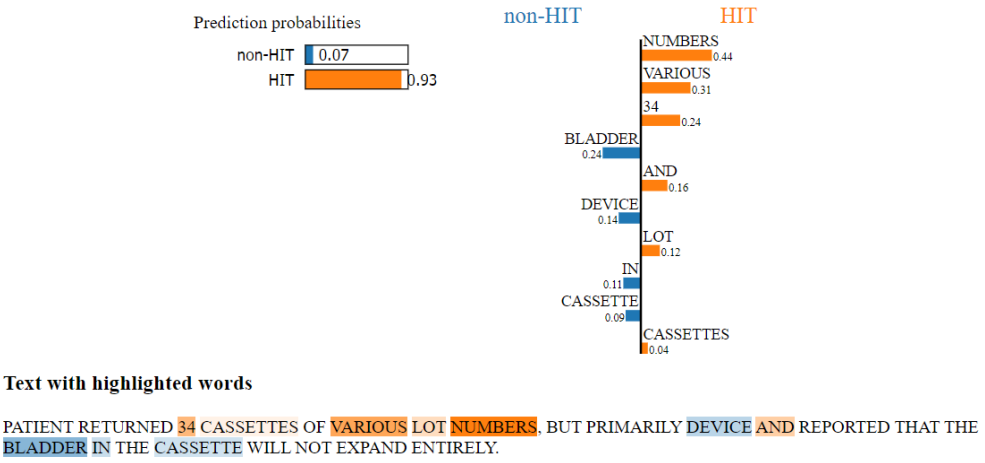


Figure 4.14: LIME applied to a false-positive classification. Features with higher weights such as “numbers”, “34”, and “various” lead the model toward the wrong classification, being the report related to a medication cassette reservoir.

4.3.3 Environmental Impact of Artificial Intelligence

As a final consideration, the environmental impact of AI should also be cited. Indeed, whenever AI models are described and analysed, a parallel discussion on the environmental impact of such technologies should be performed. CO_2 emission from the world computing infrastructure is now equivalent to aeronautics at its top, and it is growing faster and faster each year [18]. According to Belkhir and Elmeligi [9] the demand for electricity from data centres will contribute to 14% of the global emission of greenhouse gas by 2040. Even only the training of a small NLP model can produce about 300,000 kilograms of CO_2, which is the same as five gas-fuelled cars in their whole life cycle, or 125 flights from New York to Beijing and back [99]. In

such a scenario, especially when discussing the scalability of an AI model to larger case studies and for bigger stakeholders, the authors think that the common practice of maximizing computing cycles to improve performances must not be the only goal, but it has to be combined with the analysis of the consumption of energy of CPUs and GPUs, not only from the economic point of view but also as a contribution to the global pollution.

4.4 Conclusion

The proposed framework employs NLP techniques and a BERT-based model to automatically identify adverse event reports related to HITs. Input data come from the FDA MAUDE database of medical device adverse event reports, but they can also originate from different sources. The framework aims to extract RWE to support EBM of medical devices, HTA, and HTM, as well as PMS as outlined in the EU-MDR. The designed model uses a pre-trained version of Bio_ClinicalBERT, additionally fine-tuned on 2,964 adverse events reports extracted from the FDA MAUDE database, which had been previously manually labelled by experts. The model was then been tested with 741 reports. Results show better metrics than other existing NLP HIT adverse events reports text classifiers based on non-BERT models [15, 33, 63, 116]. XAI techniques have also been employed to understand how the model interprets each feature, calculating the overall weight of each word in relation to the final output classes. Both employed XAI methods (LIME and SHAP) highlight how a subset of specific features (e.g., "handheld", "computer", "software") have a high weight in determining the final output class of the model, as it is conceivable and congruent with the type of analysed events.

Highlighting both the most common faults and the unexpected new challenges before introducing a new device is vital to perform an actual assessment of the whole life-cycle of the technology (from purchase to maintenance until discontinuation), evaluating all the possible hidden costs which it may impact. The performance and the robustness of the model can be further improved by exploiting new adverse event reports extracted by the MAUDE or other SRS databases (e.g., the EU EUDAMED and the Australian DAEN). In doing so, the results that emerged from the implementation of XAI methods can be incorporated to ease the process of labelling new records.

Chapter 5

Conclusions

The primary objective of the developed framework is to extract Real-World Evidence (RWE) to support Evidence-Based Maintenance (EBM) for medical devices, Health Technology Management (HTA), Health Technology Assessment (HTM), and Post Market Surveillance (PMS) in line with the EU Medical Device Regulation 2017/745 (EU-MDR). Highlighting both common issues and unforeseen challenges before introducing a new medical device is crucial for a comprehensive assessment of the entire technology lifecycle, spanning from procurement and maintenance to eventual decommissioning. The literature review in Chapter 2 showed that the application of Natural Language Processing (NLP) on medical device adverse events (especially software) is still in the embryonic phase, so this work tries to establish a novel contribution to applying existing Deep Learning methodologies to solve unfaced clinical and management problems.

The developed framework leverages NLP techniques and Large Language Models, employing a pre-trained BERT model (namely, the Bio_ClinicalBERT model), fine-tuned using 2,964 adverse event reports extracted from the FDA MAUDE database, to automatically detect those associated with Health Information Technology (HIT). These reports had previously undergone manual expert labelling. Subsequently, the model was tested with 741 unseen reports, demonstrating superior performance metrics compared to other NLP models for adverse event reports in HIT that were not based on BERT [15, 33, 63, 116].

To enhance the transparency and interpretability of the model's decision-making process, Explainable Artificial Intelligence (XAI) techniques were utilized. These techniques, specifically LIME and SHAP, highlighted the

significance of individual features, showcasing how certain features strongly influence the model's output, aligning with the nature of the analysed events.

The model's performance and reliability can be further enhanced by leveraging new adverse event reports sourced from MAUDE or other similar Spontaneous Reporting System (SRS) databases, such as EU EUDAMED and the Australian DAEN. Additionally, insights gained from the implementation of XAI methods can be incorporated to facilitate the labelling process for new records. Further opportunities for data access may arise, including the possibility of obtaining data from local hospitals' Computerized Maintenance Management System (CMMS) software.

However, a comprehensive review of existing literature performed in Chapter 2 has revealed a notable absence of an up-to-date global standard for naming and coding medical devices and their associated fault codes in maintenance work orders. This deficiency poses significant challenges when attempting to collect data from diverse systems, as mapping across disparate nomenclatures becomes exceedingly difficult due to the unique internal organization of each nomenclature and CMMS software. These challenges hinder the extraction of harmonised RWD, which, in turn, constrains the development of EBM. This approach could otherwise offer guidance for enhancing the maintenance of medical devices while ensuring their safety and reliability.

While progress is being made in standardising the nomenclature of medical devices, largely thanks to initiatives by organisations like the World Health Organization (WHO) and the development of the International Classification and Nomenclature of Medical Devices, the standardisation of failure codes for maintenance remains in its infancy. Involved stakeholders appear more inclined to propose entirely new classifications rather than striving for interoperability with existing methodologies.

In such context, semantic ontologies emerge as valuable tools for knowledge representation, employing abstract concepts to comprehensively describe a given subject by capturing entities and their relationships. A scoping review was conducted to assess existing ontologies capable of accommodating various use cases in the healthcare environment. The review identified 32 studies, which were classified into ten areas of interest. These areas included technology, disease vocabulary, medical vocabulary, medical data, human roles, building management, drugs, services, medical procedures, and emergencies. A set of 34 ontologies extracted from these studies was also analysed and discussed in Section 3.4.

Semantic ontologies offer the potential to establish a suitable level of abstraction for sharing and reusing concepts in a standardised manner. This ensures that data from diverse sources can be provided with a common nomenclature, facilitating communication among stakeholders, and streamlining the integration of the proposed NLP framework for HTM and HTA.

Appendix A

Appendix

This appendix is related to the results of the scoping review, previously presented in Chapter 3.

Table A.1: Summary of the characteristics the selected articles.

F. Author and Year	**Title**	**Aim of the article**	**Identified ontologies**	**Semantic Area**
Cardoso 2022 [13]	DCSO: towards an ontology for machine-actionable data management plans	Presenting DMP Common Standard Ontology as a serialisation of the DCS core concepts.	DCSO	Medical Data
Van Damme 2022 [102]	Performance assessment of ontology matching systems for FAIR data	Assessing the performance of NCIT and ORDO ontologies, as well as SNOMED-CT vocabulary, as matching services for querying FAIR data.	NCIT ORDO	Disease Vocabulary
Esfahani 2022 [27]	Ontology for Symptomatic Treatment of Multiple Sclerosis	Providing an ontology for Multiple Sclerosis symptomatic treatment.	STMSO	Disease Vocabulary
Donkers 2022 [22]	Semantic Web Technologies for Indoor Environmental Quality: A Review and Ontology Design	Presenting the Building Performance Ontology	BOP	Buildings
Joo 2022 [62]	A Flexible Semantic Ontological Model Framework and Its Application to Robotic Navigation in Large Dynamic Environments	Presenting semantic navigation framework based on a Triplet Ontological Semantic Model (TOSM) to manage various conditions affecting the execution of tasks. The framework allows robots with different kinematics to perform tasks in indoor and outdoor environments.	KAON	Technology

Kim 2022 [66]	Developing a Dietary Lifestyle Ontology to Improve the Interoperability of Dietary Data: Proof-of-Concept Study	Developing a Dietary Lifestyle Ontology (DILON) and demonstrate the improved interoperability of questionnaire-based dietary data by annotating its main semantics with DILON.	DILON	Medical Vocabulary
Silva 2022 [97]	Ontologies and Knowledge Graphs in Oncology Research	Review on the last decade of works on using ontologies in cancer.	NCIT ICD-O OBI	Disease Vocabulary
Bona 2022 [12]	Semantic Integration of Multi-Modal Data and Derived Neuroimaging Results Using the Platform for Imaging in Precision Medicine (PRISM) in the Arkansas Imaging Enterprise System (ARIES)	Research data management system which features integrated capabilities to support semantic representations of multi-modal data from disparate sources across common image-processing stages, as well as derived results.	OBI NCIT IAO FMA NCBIT	Medical Vocabulary
Narayanasamy 2022 [81]	A Contemporary Review on Utilizing Semantic Web Technologies in Healthcare, Virtual Communities, and Ontology-Based Information Processing Systems	Review of using the semantic web in the domain of healthcare and virtual communities.	MedDRA PMR	Disease Vocabulary

Liu 2022 [69]	Semantic Association and Decision-Making for the Internet of Things Based on Partial Differential Fuzzy Unsupervised Models	Presents study and analysis of IoT semantic association and decision-making using a partial differential fuzzy unsupervised approach.	SSN	Technology
Cornejo-Lupa 2021 [17]	Ontoslam: An ontology for representing location and simultaneous mapping information for autonomous robots	A complete ontology to model all aspects related to autonomous robots and the Simultaneous Localization and Mapping (SLAM) problem.	CORA OntoSLAM	Technology
Babcock 2021 [5]	The Infectious Disease Ontology in the age of COVID-19	Description of IDO and its extensions integrated with the analysis of COVID-19 data: VIDO, CIDO and IDO-COVID-19.	CIDO IDO IDO-COVID-19 VIDO	Disease Vocabulary
Rahman 2021 [92]	A light-weight dynamic ontology for Internet of Things using machine learning technique	Proposal of a dynamic ontology to achieve semantic interoperability among heterogeneous devices and applications.	OneM2M IoT-Lite SSN	Technology
Maitra 2021 [74]	Using ethnographic methods to classify the human experience in medicine: A case study of the presence ontology	Creation of a conceptual framework to describe and classify data about presence, the domain of interpersonal connection in medicine.	PREO	Human Role

Mazimwe 2021 [76]	Implementation of FAIR principles for ontologies in the disaster domain: A systematic literature review	Systematic search and review of publications in the disaster management domain.	Empathi Disaster-Ontology EF SOKNOS MOAC	Emergency
Gordon 2021 [38]	People, Projects, Organizations, and Products: Designing a Knowledge Graph to Support Multi-Stakeholder Environmental Planning and Design	Developing a prototype knowledge graph based on RDF and GeoSPARQL standards for a multi-stakeholder regional environmental planning and design initiatives.	Organization-Ontology Relation-Ontology VIVO	Human Role
Bassier 2020 [8]	Processing existing building geometry for reuse as Linked Data	Combine building component information extracted from 2D images and 3D building models and publish them as Linked Data.	BOT	Buildings
Hakimi 2020 [39]	The Devices, Experimental Scaffolds, and Biomaterials Ontology (DEB): A Tool for Mapping, Annotation, and Analysis of Biomaterials Data	Description of DEB, an open resource for organizing information about biomaterials, their design, manufacture and biological testing.	DEB	Medical Vocabulary

Elsaleh 2020 [26]	IoT-Stream: A Lightweight Ontology for Internet of Things Data Streams and its Use with Data Analytics and Event Detection Services	Presentation of IoT-Stream, a lightweight instantiation of the semantic sensor network (SSN) ontology to describe the key IoT concepts that allows interoperability and discovery of sensory data in heterogeneous IoT platforms.	IoT-Stream	Technology
Yu 2019 [125]	ODAE: Ontology-based systematic representation and analysis of drug adverse events and its usage in study of adverse events given different patient age and disease conditions	Description of ODAE and its developments.	OAE ODAE	Medical Vocabulary
El-Sappagh 2018 [25]	SNOMED-CT standard ontology based on the ontology for general medical science	Development of an upper-level ontology to be used as the basis for defining the terms in SNOMED-CT.	SCTO	Medical Vocabulary
Gibaud 2018 [36]	Toward a standard ontology of surgical process models	Presentation of the OntoSPM Collaborative Action, which serves as a platform developing ontologies in the domain of surgery, focusing on surgical process modelling in the context of Surgical Data Science.	OntoSPM	Medical Procedure

Haller 2018 [40]	The modular SSN ontology: A joint W3C and OGC standard specifying the semantics of sensors, observations, sampling, and actuation	Overview of the SSN, SOSA and SSNZ ontologies discussing the rationale behind key design decisions and main differences.	SSN SOSA SSNX	Technology
Glockner 2017 [37]	Lose ODP - an ontology design pattern for logistics services	Presentation of an ontology design pattern for logistics services	LoSe ODP	Services
Hicks 2016 [44]	The ontology of medically related social entities: Recent developments	Description of OMRSE and its recent developments.	OMRSE	Human Role
Santana 2015 [95]	Towards a Formal Representation of Processes and Objects Regarding the Delivery of Telehealth Services: The Telehealth Ontology (TEON)	Description and Presentation of TEON, elucidating its main use-case, its applicability and potential to improve information exchange, interoperability and decision support.	TEON	Technology
McMurray 2015 [77]	Ontological modelling of electronic health information exchange	Description of the conceptual framework of health system information exchange and its related ontology.	HEIO	Medical Data
Hanna 2013 [41]	Building a drug ontology based on RxNorm and other sources	Description of the building and the structure of the Drug Ontology.	DrOn RxNorm	Drugs

Prestes 2013 [90]	Towards a core ontology for robotics and automation	Presentation of the current results of the newly formed IEEE-RAS Working Group, named Ontologies for Robotics and Automation and introduction of a core ontology that encompasses a set of terms commonly used in Robotics and Automation.	CORA	Technology
Ison 2013 [59]	EDAM: An ontology of bioinformatics operations, types of data and identifiers, topics and formats	Presenting EDAM, an ontology of bioinformatics operations, types of data and identifiers, data formats and topics with the goal of creating machine-understandable annotations for use within resource catalogues.	EDAM	Medical Data
Liu 2011 [68]	Effectiveness of lexico-syntactic pattern matching for ontology enrichment with clinical documents	Evaluate the effectiveness of a Lexico-Syntactic Pattern matching method for ontology enrichment using clinical documents.	NCIT RadLex	Medical Vocabulary
Robinson 2010 [94]	The Human Phenotype Ontology	Development and description of HPO to capture phenotypic information.	HPO	Disease Vocabulary

Table A.2: Table of identified ontologies

Ontology	Description	IRI	Website	Topic	Ref.
BOP	The Building Performance Ontology aims to enable the integration of topological building information with static and dynamic properties, to create a homogeneous data environment used by complex building performance assessments.	https://alexdonkers.github.io/bop/index.html	https://alexdonkers.github.io/bop/index.html	Infrastructures	[22]
BOT	The Building Topology Ontology is a minimal ontology for describing the core topological concepts of a building.	https://w3c-lbd-cg.github.io/bot	https://w3c-lbd-cg.github.io/bot	Infrastructures	[8]
CIDO	CIDO aims to ontologically represent and standardize various aspects of coronavirus infectious diseases, including their etiology, transmission, epidemiology, pathogenesis, diagnosis, prevention and treatment.	http://purl.obolibrary.org/obo/cido.owl	https://github.com/cido-ontology/cido	COVID-19	[5]
CORA	Core Ontology for Robotics and Automation	http://purl.org/ieee1872-owl/cora	https://github.com/srfiorini/IEEE1872-owl	Robotics	[17, 90]

DCSO	The DMP Common Standard Ontology aims to represent the DMP Common Standard model, through the usage of semantic web technology. It represents the DMP Common Standard model using the Web Ontology Language (OWL).	http://semantics.id/ns/dcso/index-en.html	http://semantics.id/ns/dcso/index-en.html	Data Management	[13]
DEB	Devices, Experimental scaffolds and Biomaterials Ontology: it is an ontology built to facilitate data cataloguing in the field of medical devices, experimental scaffolds and biomaterials.	http://www.semanticweb.org/osnathakimi/ontologies/deb	https://projectdebbie.github.io	Biomaterials	[39]
DILON	The Dietary Lifestyle Ontology (DILON) aims to represent dietary lifestyle data. Concepts are pulled from Korean dietary assessment scales and English assessment scales. Concepts are labeled in English and Korean translations are also provided.	https://bioportal.bioontology.org/ontologies/DILON	https://bioportal.bioontology.org/ontologies/DILON	Dietary Lifestyle Data	[66]

Disaster Ontology	The Disaster ontology is an ontology for disaster control and it captures all the entities concerning disasters.	http://www.semanticweb.org/ontologies/2008/10/Ontology1226057991156.owl	https://onki.fi/en/browser/	Emergency Management	[76]
DRON	The Drug Ontology: it is the deposit for the Drug Ontology, an ontology of drug products, their ingredients, and their packaging.	http://purl.obolibrary.org/obo/dron.owl	https://bitbucket.org/uamsdbmi/dron/src/master	Drugs	[41]
EDAM	Bioscientific data analysis ontology and data management: EDAM is a complete ontology of well-established, entities that are widespread within computational biology and bioinformatics.	http://edamontology.org	http://edamontology.org/page	Informatics, Workflow Management	[59]
Empathi	Ontology that conceptualizes the core concepts concerning the domain of emergency managing and planning of hazard crises.	https://w3id.org/empathi/	https://shekarpour.github.io/empathi.io/	Emergency Management	[76]
HEIO	Regional Healthcare System Interoperability and Information Exchange Measurement Ontology.	http://whistl.uwaterloo.ca/heio.owl	https://bioportal.bioontology.org/ontologies/HEIO	Electronic Health Records	[77]

HPO	Human Phenotype Ontology: it provides a standardized vocabulary of phenotypic abnormalities and clinical features encountered in human disease.	http://purl.obolibrary.org/obo/hp.owl	https://hpo.jax.org/app/	Diseases	[94]
ICDO	International Classiffication of Disease Ontology: it is the ontology representation of the ICD system.	http://purl.obolibrary.org/obo/icdo/	https://github.com/icdo/ICDO	Diseases	[97]
IDO Core	Infectious Disease Ontology	http://purl.obolibrary.org/obo/ido.owl	https://www.bioontology.org/wiki/Infectious_Disease_Ontology	Infectious Diseases	[5]
IDO-COVID-19	The COVID-19 Infectious Disease Ontology: it is an extension of the Infectious Disease Ontology (IDO) and the Virus Infectious Disease Ontology (VIDO)	http://purl.obolibrary.org/obo/ido-covid-19.owl	https://www.ebi.ac.uk/ols/ontologies/idocovid19	COVID-19	[5]
LoSe ODP	Logistics Service Ontology Design Pattern: it is the ontology design pattern about logistics services.	https://github.com/Michael-Gloeckner/LoSeOn	https://github.com/Michael-Gloeckner/LoSe_ODP	Logistics	[37]

NCIT	NCI Thesaurus: it is a reference terminology that includes broad coverage of the cancer domain, including cancer related diseases, findings and abnormalities.	http://purl.obolibrary.org/obo/ncit.owl	https://github.com/NCI-Thesaurus/thesaurus-obo-edition	Human Diseases, Clinical Terminology	[12, 68,97, 102]
OAE	Ontology of Adverse Events: it is developed to standardize adverse event annotation, integrate various adverse event data, and support computer-assisted reasoning.	http://purl.obolibrary.org/obo/oae.owl	http://www.oae-ontology.org/	Adverse Events	[44, 125]
OBI	Ontology for Biomedical Investigations: an integrated ontology for the description of life-science and clinical investigations.	http://purl.obolibrary.org/obo/obi.owl	http://obi-ontology.org/	Data Collection	[12, 97]
ODAE	Ontology of Drug Adverse Events: biomedical ontology in the area of drug adverse events.	https://raw.githubusercontent.com/ODAE-ontology/ODAE/master/src/ontology/odae_merged.owl	https://github.com/ODAE-ontology/ODAE	Adverse Events and Drugs	[125]

OGMS	Ontology for General Medical Science: ontology of entities involved in a clinical encounter. OGMS includes very general terms that are used across medical disciplines, including: 'disease', 'disorder', 'disease course', 'diagnosis', 'patient', and 'healthcare provider'.	http://purl.obolibrary.org/obo/ogms.owl	https://github.com/OGMS/ogms	Human Diseases, Clinical Terminology	[5,25]
OMRSE	Ontology of Medically Related Social Entities: this ontology covers the domain of social entities that are related to health care, such as demographic information and the roles of various individuals and organizations.	http://purl.obolibrary.org/obo/omrse.owl	https://github.com/ufbmi/OMRSE/wiki/OMRSE-Overview	Clinical Staff	[44]
OneM2M	OneM2M's Base Ontology constitutes a basis framework for specifying the semantics of data that are handled in oneM2M. Sub-classes of some of its concepts are expected to be defined by other bodies in order to enable semantic interworking.	https://git.onem2m.org/MAS/BaseOntology/raw/master/base_ontology.owl	https://www.onem2m.org/technical/onem2m-ontologies	IoT	[92]

OntoSLAM	The ontology models all aspects related to autonomous robots and the SLAM problem.	https://github.com/Alex23013/ontoSLAM	https://github.com/Alex23013/ontoSLAM	Robots	[17]
OntoSPM	OntoSPM, a core ontology for surgical process models: motivations, working assumptions and current status.	http://medicis/spm.owl/OntoSPM	https://ontospm.univ-rennes1.fr/doku.php?id=ontology	Surgery	[36]
Organization Ontology	This ontology is designed to enable publication of information on organizations and organizational structures including governmental organizations.	http://www.w3.org/ns/org	https://www.w3.org/TR/vocab-org/	Clinical Staff	[38]
PMR	Physical Medicine and Rehabilitation: Knowledge representation related to computer-based decision support in rehabilitation.	http://purl.bioontology.org/ontology/PMR.owl	https://bioportal.bioontology.org/ontologies/PMR	Rehabilitation	[81]

PREO	Presence Ontology: it is a systematic vocabulary of terms with defined relationships that models the encounters taking place every day among providers, patients, and family members or friends in environments such as hospitals and clinics. The Presence Ontology provides a conceptual model for the human experience in medicine.	http://presence-ontology.org/ontology/	https://med.stanford.edu/presence/about.html	Clinical Staff	[74]
SCTO	Systematized Nomenclature of Medicine Clinical Terms Ontology	https://bioportal.bioontology.org/ontologies/SCTO	https://bioportal.bioontology.org/ontologies/SCTO	Clinical Terms	[25, 97]
SSN	The Semantic Sensor Network (SSN) ontology is an ontology for describing sensors and their observations, the involved procedures, the studied features of interest, the samples used to do so, and the observed properties, as well as actuators.	https://www.w3.org/TR/vocab-ssn	https://www.w3.org/TR/vocab-ssn	Sensors	[40, 69,92]

STMSO	The Symptomatic Treatment of Multiple Sclerosis Ontology is the first comprehensive semantic representation of symptomatic treatment of MS and provides a major step toward the development of intelligent Clinical Decision Support System (CDSS) for MS symptomatic treatment.	https://bioportal.bioontology.org/ontologies/STMSO	https://bioportal.bioontology.org/ontologies/STMSO	Diseases	[27]
TEON	Telehealth Ontology	http://www.nutes.ufpe.br/teon.owl	https://github.com/TelehealthOntology/teon	Telehealth	[95]
VIDO	The Virus Infectious Disease Ontology:The Virus Infectious Disease Ontology (IDO Virus) is an extension of the Infectious Disease Ontology (IDO).	http://purl.obolibrary.org/obo/vido.owl	https://www.ebi.ac.uk/ols/ontologies/vido	Diseases	[5]

Appendix B

Publications

This research activity has led to several publications in international journals and conferences. These are summarized below.[1]

International Journals

1. **A. Luschi**, P. Nesi, and E. Iadanza, "Evidence-based Clinical Engineering: Health Information Technology Adverse Events Identification and Classification with Natural Language Processing", *Heliyon*, vol. 9(11), e21723, 2023, [DOI:10.1016/j.heliyon.2023.e21723].

2. **A. Luschi**, C. Petraccone, G. Fico, L. Pecchia, and E. Iadanza, "Semantic Ontologies for Complex Healthcare Structures: A Scoping Review", *IEEE Access*, vol. 11, pp. 19228-19246, 2023, [DOI:10.1109/ACCESS.2023.3248969].

International Conferences and Workshops

1. E. Iadanza and **A. Luschi**, "Standardization of failure codes and nomenclature of medical devices for Evidence-Based Maintenance", in *IFMBE Proceedings*, 16th Mediterranean Conference on Medical and Biological Engineering and Computing (MEDICON) and 5th International Conference on Medical and Biological Engineering (CMBEBIH), Sarajevo (Bosnia and Herzegovina), vol. 94, pp. 170-177, 2024.

2. **A. Luschi**, P. Nesi, and E. Iadanza, "Evidence Based Management of medical devices using natural language processing and neural networks to study

[1] The author's bibliometric indices are the following: H-index = 9, total number of citations = 167 (source: Scopus on Month 1, 2024).

medical devices failures", in *IFMBE Proceedings*, IUPESM World Congress On Medical Physics and Biomedical Engineering, Singapore (Singapore), vol. In Press, 2022.

3. L. Mascii, **A. Luschi**, and E. Iadanza, "Sentiment Analysis for Performance Evaluation of Maintenance in Healthcare", in *IFMBE Proceedings*, 4th International Conference on Medical and Biological Engineering (CMBEBIH), Mostar (Bosnia and Herzegovina), vol. 84, pp. 359-367, 2021, [DOI:10.1007/978-3-030-73909-6_41]. **(IFMBE Best 2022 manuscript award)**

National Conferences

1. **A. Luschi**, P. Nesi, and E. Iadanza, "Health Information Technology Adverse Events Identification and Classification with Natural Language Processing and Deep Learning", in *Gruppo Nazionale Bioingegneria (GNB)*, 8th National Congress of Bioengineering, Padova (PD), Italy, pp. 142-145, 2023.

2. **A. Luschi**, P. Nesi, and E. Iadanza, "Evidence Based Management: intelligenza artificiale a supporto della classificazione degli eventi avversi per le tecnologie informatiche sanitarie", in *XXIII Convegno Nazionale Associazione Italiana Ingegneri Clinici*, Firenze (FI), Italy, 2023.

Bibliography

[1] A. Akkasi and M.-F. Moens, "Causal relationship extraction from biomedical text using deep neural models: A comprehensive survey," *Journal of biomedical informatics*, p. 103820, 2021.

[2] H. Alemzadeh, R. K. Iyer, Z. Kalbarczyk, and J. Raman, "Analysis of safety-critical computer failures in medical devices," *IEEE Security Privacy*, vol. 11, no. 4, pp. 14–26, 2013.

[3] E. Alsentzer, J. Murphy, W. Boag, W.-H. Weng, D. Jindi, T. Naumann, and M. McDermott, "Publicly available clinical BERT embeddings," in *Proceedings of the 2nd Clinical Natural Language Processing Workshop*. Minneapolis, Minnesota, USA: Association for Computational Linguistics, 2019, pp. 72–78. [Online]. Available: https://aclanthology.org/W19-1909

[4] R. Arp, B. Smith, and A. D. Spear, *Building ontologies with Basic Formal Ontology*. Cambridge: MIT Press, 2015.

[5] S. Babcock, J. Beverley, L. G. Cowell, and B. Smith, "The infectious disease ontology in the age of covid-19," *Journal of biomedical semantics*, vol. 12, no. 1, p. 13, 2021.

[6] M. F. Baretich and C. Davis-Smith, "Optimizing the cmms failure code field," Association for the Advancement of Medical Instrumentation, Tech. Rep., 2020.

[7] S. Barocas, S. Friedler, M. Hardt, J. Kroll, S. Venka-Tasubramanian, and H. Wallach, "The fat-ml workshop series on fairness, accountability, and transparency in machine learning," *cit. on*, p. 7, 2018.

[8] M. Bassier, M. Bonduel, J. Derdaele, and M. Vergauwen, "Processing existing building geometry for reuse as linked data," *Automation in Construction*, vol. 115, p. 103180, 2020.

[9] L. Belkhir and A. Elmeligi, "Assessing ict global emissions footprint: Trends to 2040 & recommendations," *Journal of Cleaner Production*, vol. 177, pp. 448–463, 2018.

[10] I. Beltagy, K. Lo, and A. Cohan, "SciBERT: A Pretrained Language Model for Scientific Text," in *EMNLP*, 2019.

[11] Z. Bliznakov, G. Mitalas, and N. Pallikarakis, "Analysis and classification of medical device recalls," in *World Congress on Medical Physics and Biomedical Engineering 2006*, R. Magjarevic and J. H. Nagel, Eds. Springer Berlin Heidelberg, 2007, pp. 3782–3785.

[12] J. Bona, A. Kemp, C. Cox, T. Nolan, L. Pillai, A. Das, J. Galvin, L. Larson-Prior, T. Virmani, and F. Prior, "Semantic integration of multi-modal data and derived neuroimaging results using the platform for imaging in precision medicine (prism) in the arkansas imaging enterprise system (aries)," *Frontiers in Artificial Intelligence*, vol. 4, 2022.

[13] J. Cardoso, L. Castro, F. Ekaputra, M. Jacquemot, M. SuchÃ¡nek, T. Miksa, and J. Borbinha, "Dcso: Towards an ontology for machine-actionable data management plans," *Journal of Biomedical Semantics*, vol. 13, no. 1, 2022.

[14] A. Cave, X. Kurz, and P. Arlett, "Real world data for regulatory decision making: Challenges and possible solutions for europe," *Clinical Pharmacology & Therapeutics*, vol. 106, 2019.

[15] K. Chai, S. Anthony, E. Coiera, and F. Magrabi, "Using statistical text classification to identify health information technology incidents," *Journal of the American Medical Informatics Association : JAMIA*, vol. 20, 2013.

[16] G. Chung, K. Etter, and A. Yoo, "Medical device active surveillance of spontaneous reports: A literature review of signal detection methods," *Pharmacoepidemiol Drug Saf*, vol. 29, no. 4, pp. 369–379, 2020.

[17] M. A. Cornejo-Lupa, Y. Cardinale, R. Ticona-Herrera, D. Barrios-Aranibar, M. Andrade, and J. Diaz-Amado, "Ontoslam: An ontology for representing location and simultaneous mapping information for autonomous robots," *Robotics*, vol. 10, no. 4, 2021.

[18] K. Crawford, *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press, 2021. [Online]. Available: http://www.jstor.org/stable/j.ctv1ghv45t

[19] K. Denecke and Y. Deng, "Sentiment analysis in medical settings: New opportunities and challenges," *Artificial Intelligence in Medicine*, vol. 64, no. 1, pp. 17–27, 2015.

[20] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "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, Volume 1 (Long and Short Papers)*. Minneapolis, Minnesota: Association for Computational Linguistics, 2019, pp. 4171–4186.

[21] M. Diaz-Galiano, M. Martin-Valdivia, and L. Urena-Lopez, "Query expansion with a medical ontology to improve a multimodal information retrieval system," *Computers in Biology and Medicine*, vol. 39, no. 4, pp. 396–403, 2009.

[22] A. Donkers, D. Yang, B. de Vries, and N. Baken, "Semantic web technologies for indoor environmental quality: A review and ontology design," *Buildings*, vol. 12, no. 10, 2022. [Online]. Available: https://www.mdpi.com/2075-5309/12/10/1522

[23] J. Du, J. Xu, H. yi Song, X. Liu, and C. Tao, "Optimization on machine learning based approaches for sentiment analysis on hpv vaccines related tweets," *Journal of Biomedical Semantics*, vol. 8, 2017.

[24] S. El Anigri, M. M. Himmi, and A. Mahmoudi, "How BERT's Dropout Fine-Tuning Affects Text Classification?" in *Business Intelligence*, M. Fakir, M. Baslam, and R. El Ayachi, Eds. Springer International Publishing, 2021, pp. 130–139.

[25] S. El-Sappagh, F. Franda, F. Ali, and K. Kwak, "Snomed ct standard ontology based on the ontology for general medical science," *BMC Medical Informatics and Decision Making*, vol. 18, 2018.

[26] T. Elsaleh, S. Enshaeifar, R. Rezvani, S. T. Acton, V. Janeiko, and M. Bermudez-Edo, "Iot-stream: A lightweight ontology for internet of things data streams and its use with data analytics and event detection services," *Sensors*, vol. 20, no. 4, 2020.

[27] A. I. Esfahani Misagh Zahiri, Ahmadi Maryam, "Ontology for symptomatic treatment of multiple sclerosis," *Healthc Inform Res*, vol. 28, no. 4, pp. 332–342, 2022.

[28] European Commission, "Leveraging ai based technology to transform the future of health care delivery in leading hospitals in europe," https://cordis.europa.eu/project/id/101017331, 2021, accessed: 2023-10-27.

[29] European Union, "Medical devices regulation (eu) 2017/745," https://eur-lex.europa.eu/legal-content/IT/ALL/?uri=celex:32017R0745, 2017, accessed: 2023-03-23.

[30] B. Farahani, F. Firouzi, and K. Chakrabarty, *Healthcare IoT*. Cham: Springer International Publishing, 2020, pp. 515–545.

[31] B. Farahani, F. Firouzi, V. Chang, M. Badaroglu, N. Constant, and K. Mankodiya, "Towards fog-driven iot ehealth: Promises and challenges of iot in medicine and healthcare," *Future Generation Computer Systems*, vol. 78, pp. 659–676, 2018.

[32] F. Firouzi, A. M. Rahmani, K. Mankodiya, M. Badaroglu, G. Merrett, P. Wong, and B. Farahani, "Internet-of-things and big data for smarter healthcare: From device to architecture, applications and analytics," *Future Generation Computer Systems*, vol. 78, pp. 583–586, 2018.

[33] A. Fong, K. Adams, M. Gaunt, J. Howe, K. Kellogg, and R. Ratwani, "Identifying health information technology related safety event reports from patient safety event report databases," *Journal of Biomedical Informatics*, vol. 86, 2018.

[34] O. B. O. Foundry, "Ontobee," https://ontobee.org/, 2022, accessed: 2022-07-11.

[35] A. Gangemi and V. Presutti, *Ontology Design Patterns*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 221–243.

[36] B. Gibaud, G. Forestier, C. Feldmann, G. Ferrigno, P. GonÃ§alves, T. Haidegger, C. Julliard, D. Katic, H. Kenngott, L. Maier-Hein, K. MÃ¤rz, E. De Momi, D. Nagy, H. Nakawala, J. Neumann, T. Neumuth, J. Balderrama, S. Speidel, M. Wagner, and P. Jannin, "Toward a standard ontology of surgical process models," *International Journal of Computer Assisted Radiology and Surgery*, vol. 13, 2018.

[37] M. Glockner and A. Ludwig, *Lose ODP - an ontology design pattern for logistics services*. IOS Press, 2017, vol. 32, pp. 131–143.

[38] S. N. Gordon, P. J. Murphy, J. A. Gallo, P. Huber, A. Hollander, A. Edwards, and P. Jankowski, "People, projects, organizations, and products:

Designing a knowledge graph to support multi-stakeholder environmental planning and design," *ISPRS International Journal of Geo-Information*, vol. 10, no. 12, 2021.

[39] O. Hakimi, J. Gelpi, M. Krallinger, F. Curi, D. Repchevsky, and M.-P. Ginebra, "The devices, experimental scaffolds, and biomaterials ontology (deb): A tool for mapping, annotation, and analysis of biomaterials' data," *Advanced Functional Materials*, p. 13, 2020.

[40] A. Haller, K. Janowicz, S. Cox, M. LefranÃ§ois, K. Taylor, D. Phuoc, J. Lieberman, R. GarcÃa Castro, R. Atkinson, and C. Stadler, "The modular ssn ontology: A joint w3c and ogc standard specifying the semantics of sensors, observations, sampling, and actuation," *Semantic Web*, vol. 10, 2018.

[41] J. Hanna, E. Joseph, M. Brochhausen, and W. Hogan, "Building a drug ontology based on rxnorm and other sources," *Journal of biomedical semantics*, vol. 4, p. 44, 2013.

[42] R. Harpaz, S. Vilar, W. DuMouchel, H. Salmasian, K. Haerian, N. Shah, and H. Chase, "Combing signals from spontaneous reports and electronic health records for detection of adverse drug reactions," *Journal of the American Medical Informatics Association : JAMIA*, vol. 20, 2012.

[43] Y. He, S. Sarntivijai, Y. Lin, Z. Xiang, A. Guo, S. Zhang, D. Jagannathan, L. Toldo, C. Tao, and B. Smith, "Oae: the ontology of adverse events," *Journal of biomedical semantics*, vol. 5, no. 1, pp. 1–13, 2014.

[44] A. Hicks, J. Hanna, D. Welch, M. Brochhausen, and W. Hogan, "The ontology of medically related social entities: Recent developments," *Journal of Biomedical Semantics*, vol. 7, 2016.

[45] HL7, "HL7 FHIR - Release 4B," http://www.hl7.org/fhir/rdf.html#ontologie, 2022, accessed: 2023-01-18.

[46] K. Huang, J. Altosaar, and R. Ranganath, "ClinicalBERT: Modeling Clinical Notes and Predicting Hospital Readmission," *ArXiv*, vol. abs/1904.05342, 2019.

[47] E. Iadanza, S. Cerofolini, C. Lombardo, F. Satta, and M. Gherardelli, "Medical devices nomenclature systems: a scoping review," *Health and Technology*, vol. 11, pp. 1–12, 2021.

[48] E. Iadanza, F. Dori, R. Miniati, and E. Corrado, "Electromagnetic interferences (emi) from active rfid on critical care equipment," in *IFMBE Proceedings*, vol. 29, 2010, pp. 991–994.

[49] E. Iadanza, V. Gonnelli, F. Satta, and M. Gherardelli, "Evidence-based medical equipment management: a convenient implementation," *Medical & Biological Engineering & Computing*, vol. 57, pp. 1–16, 2019.

[50] E. Iadanza, C. Ignesti, R. Miniati, and A. Luschi, "Risk management process on a new microwave thermal ablation device: Assessment and follow up," in *XIV Mediterranean Conference on Medical and Biological Engineering and Computing 2016*, vol. 57, 2016, pp. 1019–1023.

[51] IEEE Standards Association, "Core ontology for robotics and automation," https://github.com/srfiorini/IEEE1872-owl, 2015, accessed: 2022-07-12.

[52] Institute of Electrical and Electronics Engineers, IEEE Robotics and Automation Society. Standing Committee for Standards Activities, "1872-2015 - ieee standard ontologies for robotics and automation," Institute of Electrical and Electronics Engineers, IEEE Robotics and Automation Society. Standing Committee for Standards Activities, Standard, 2015.

[53] International Medical Device Regulators Forum, "IMDRF Terminologies for Categorized Adverse Event Reporting (AER): Terms, Terminology Structure and Codes," https://www.imdrf.org/documents/terminologies-categorized-adverse-event-reporting-aer-terms-terminology-and-codes, 2020, accessed: 2023-03-23.

[54] International Network of Agencies for Health Technology Assessment and Health Technology Assessment international, "Health technology assessment," http://htaglossary.net/health-technology-assessment, 2022, accessed: 2022-02-08.

[55] International Organization for Standardization, "ISO 8373 Robots and robotic devices - Vocabulary," International Organization for Standardization, Standard, 2012.

[56] ——, "Building construction - Organization of information about construction works - Part 2: Framework for classification," International Organization for Standardization, Standard, 2013.

[57] ——, "ISO 15225:2016. Medical devices. Quality management. Medical device nomenclature data structure ," International Organization for Standardization, Standard, 2016.

[58] ——, "ISO/IEC 21838-2 Information technology - Top-level ontologies (TLO) - Part 2: Basic Formal Ontology (BFO)," International Organization for Standardization, Standard, 2021.

[59] J. Ison, M. KalaÅ¡, I. Jonassen, D. Bolser, M. Uludag, H. McWilliam, J. Malone, R. Lopez, S. Pettifer, and P. Rice, "EDAM: an ontology of bioinformatics operations, types of data and identifiers, topics and formats," *Bioinformatics*, vol. 29, no. 10, pp. 1325–1332, 2013.

[60] S. Iyer, R. Harpaz, P. LePendu, A. Bauer-Mehren, and N. Shah, "Mining clinical text for signals of adverse drug-drug interactions," *Journal of the American Medical Informatics Association*, vol. 21, 2013.

[61] S. M. Jimenez-Zafra, M. T. Martin-Valdivia, M. D. Molina-Gonzalez, and L. A. Urena-Lopez, "How do we talk about doctors and drugs? sentiment analysis in forums expressing opinions for medical domain," *Artificial Intelligence in Medicine*, vol. 93, pp. 50–57, 2019, extracting and Processing of Rich Semantics from Medical Texts.

[62] S. Joo, S. Bae, J. Choi, H. Park, S. Lee, S. You, T. Uhm, J. Moon, and T. Kuc, "A flexible semantic ontological model framework and its application to robotic navigation in large dynamic environments," *Electronics*, vol. 11, no. 15, 2022. [Online]. Available: https://www.mdpi.com/2079-9292/11/15/2420

[63] H. Kang and Y. Gong, "Creating a database for health it events via a hybrid deep learning model," *Journal of Biomedical Informatics*, vol. 110, p. 103556, 2020.

[64] P. Kataria, R. Juric, S. Paurobally, and K. Madani, "Implementation of ontology for intelligent hospital wards," in *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS 2008)*, 2008, pp. 253–253.

[65] S. Khan, M. Naseer, M. Hayat, S. W. Zamir, F. S. Khan, and M. Shah, "Transformers in vision: A survey," *ACM Comput. Surv.*, vol. 54, no. 10s, 2022.

[66] H. Kim, J. Jung, and J. Choi, "Developing a dietary lifestyle ontology to improve the interoperability of dietary data: Proof-of-concept study," *JMIR Form Res*, vol. 6, no. 4, p. e34962, 2022.

[67] J. Lee, W. Yoon, S. Kim, D. Kim, S. Kim, C. H. So, and J. Kang, "BioBERT: a pre-trained biomedical language representation model for biomedical text mining," *Bioinformatics*, vol. 36, pp. 1234 – 1240, 2020.

[68] K. Liu, W. Chapman, G. Savova, C. Chute, N. Sioutos, and R. Crowley, "Effectiveness of lexico-syntactic pattern matching for ontology enrichment with clinical documents," *Methods of information in medicine*, vol. 50, pp. 397–407, 2011.

[69] W. Liu and B. Lu, "Semantic association and decision-making for the internet of things based on partial differential fuzzy unsupervised models," *Mathematical Problems in Engineering*, vol. 2022, pp. 1–11, 2022.

[70] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems 30*, 2017, pp. 4765–4774.

[71] X. Luo, G. Zimet, and S. Shah, "A natural language processing framework to analyse the opinions on hpv vaccination reflected in twitter over 10 years (2008 - 2017)," *Human Vaccines & Immunotherapeutics*, vol. 15, no. 7-8, pp. 1496–1504, 2019.

[72] X. Lv, Y. Guan, J. Yang, and J. Wu, "Clinical relation extraction with deep learning," *International Journal of Hybrid Information Technology*, vol. 9, pp. 237–248, 2016.

[73] F. Magrabi, M.-S. Ong, W. B. Runciman, and E. W. Coiera, "Patient safety problems associated with heathcare information technology: an analysis of adverse events reported to the us food and drug administration." *AMIA ... Annual Symposium proceedings. AMIA Symposium*, vol. 2011, pp. 853–7, 2011.

[74] A. Maitra, M. R. Kamdar, D. M. Zulman, M. C. Haverfield, C. Brown-Johnson, R. Schwartz, S. T. Israni, A. Verghese, and M. A. Musen, "Using ethnographic methods to classify the human experience in medicine: a case study of the presence ontology," *Journal of the American Medical Informatics Association*, vol. 28, no. 9, pp. 1900–1909, 2021.

[75] L. Massai, P. Nesi, and G. Pantaleo, "Paval: A location-aware virtual personal assistant for retrieving geolocated points of interest and location-based services," *Engineering Applications of Artificial Intelligence*, vol. 77, pp. 70–85, 2019.

[76] A. Mazimwe, I. Hammouda, and A. Gidudu, "Implementation of fair principles for ontologies in the disaster domain: A systematic literature review," *ISPRS International Journal of Geo-Information*, vol. 10, no. 5, 2021.

[77] J. McMurray, L. Zhu, I. McKillop, and H. Chen, "Ontological modeling of electronic health information exchange," *Journal of Biomedical Informatics*, vol. 56, pp. 169–178, 2015.

[78] R. Miotto, L. Li, and B. Kidd, "Deep patient: An unsupervised representation to predict the future of patients from the electronic health records," *Scientific Reports*, vol. 6, p. 26094, 2016.

[79] M. Mosbach, M. Andriushchenko, and D. Klakow, "On the Stability of Fine-tuning BERT: Misconceptions, Explanations, and Strong Baselines," in *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net, 2021. [Online]. Available: https://openreview.net/forum?id=nzpLWnVAyah

[80] A. W. Muzaffar, F. Azam, and U. Qamar, "A relation extraction framework for biomedical text using hybrid feature set," *Computational and Mathematical Methods in Medicine*, vol. 2015, 2015.

[81] S. K. Narayanasamy, K. Srinivasan, Y.-C. Hu, S. K. Masilamani, and K.-Y. Huang, "A contemporary review on utilizing semantic web technologies in healthcare, virtual communities, and ontology-based information processing systems," *Electronics*, vol. 11, no. 3, 2022.

[82] National Center for Biomedical Ontology, "Bioportal," https://bioportal.bioontology.org/, 2022, accessed: 2022-07-11.

[83] M. T. Neame, G. Sefton, M. Roberts, D. Harkness, I. P. Sinha, and D. B. Hawcutt, "Evaluating health information technologies: A systematic review of framework recommendations," *International Journal of Medical Informatics*, vol. 142, p. 104247, 2020.

[84] P. Nesi, G. Pantaleo, and G. Sanesi, "A distributed framework for nlp-based keyword and keyphrase extraction from web pages and documents," in *Proceedings of the 21st Int. Conf. on Distributed Multimedia Systems (DMS'2015)*, 2015, pp. 155–161.

[85] ——, "A hadoop based platform for natural language processing of web pages and documents," *Journal of Visual Languages & Computing*, vol. 31, pp. 130–138, 2015.

[86] Odin, "Odin - smart hospitals," https://odin-smarthospitals.eu/, 2021, accessed: 2023-10-27.

[87] OntoSPM Collaborative Action, "Ontospm collaborative action," https://ontospm.univ-rennes1.fr/doku.php, 2022, accessed: 2022-07-11.

[88] L. Pecchia, D. Piaggio, N. Pallikarakis, and E. Iadanza, "Health technology assessment teaching for bme," in *Clinical Engineering Handbook (Second Edition)*, second edition ed., E. Iadanza, Ed. Academic Press, 2020, pp. 832–835.

[89] J. Polisena, M. Andellini, P. Salerno, S. Borsci, L. Pecchia, and E. Iadanza, "Case studies on the use of sentiment analysis to assess the effectiveness and safety of health technologies: A scoping review," *IEEE Access*, vol. 9, pp. 66 043–66 051, 2021.

[90] E. Prestes, J. L. Carbonera, S. Rama Fiorini, V. A. M. Jorge, M. Abel, R. Madhavan, A. Locoro, P. Goncalves, M. E. Barreto, M. Habib, A. Chibani, S. Gerard, Y. Amirat, and C. Schlenoff, "Towards a core ontology for robotics and automation," *Robotics and Autonomous Systems*, vol. 61, no. 11, pp. 1193–1204, 2013, ubiquitous Robotics.

[91] A. Radford, K. Narasimhan, T. Salimans, and I. Sutskever, "Improving language understanding by generative pre-training," *Technical report, OpenAI*, 2018.

[92] H. Rahman and M. I. Hussain, "A light-weight dynamic ontology for internet of things using machine learning technique," *ICT Express*, vol. 7, no. 3, pp. 355–360, 2021.

[93] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why Should I Trust You?: Explaining the Predictions of Any Classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, August 13-17, 2016*, 2016, pp. 1135–1144.

[94] P. Robinson and S. Mundlos, "The human phenotype ontology," *Clinical genetics*, vol. 77, pp. 525–34, 2010.

[95] F. Santana da Silva, S. Schulz, A. Campos, and M. Novaes, "Towards a formal representation of processes and objects regarding the delivery of telehealth services: The telehealth ontology (teon)," *Studies in health technology and informatics*, vol. 216, p. 1108, 2015.

[96] B. Shickel, P. J. Tighe, A. Bihorac, and P. Rashidi, "Deep EHR: A Survey of Recent Advances in Deep Learning Techniques for Electronic Health Record (EHR) Analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 22, no. 5, pp. 1589–1604, 2018.

[97] M. C. Silva, P. Eugenio, D. Faria, and C. Pesquita, "Ontologies and knowledge graphs in oncology research," *Cancers*, vol. 14, no. 8, 2022.

[98] SNOMED International, "Systematized nomenclature of medicine clinical terms," https://www.snomed.org/snomed-ct/five-step-briefing, 2021, accessed: 2022-07-12.

[99] E. Strubell, A. Ganesh, and A. McCallum, "Energy and policy considerations for deep learning in nlp," *ArXiv*, vol. abs/1906.02243, 2019.

[100] J. Sun, J. Pan, Y. Jin, Q. Zhang, Y. Lv, and J. Feng, "Establishment of a medical device adverse event management system for hospitals," *BMC health services research*, vol. 22, p. 1406, 2022.

[101] A. C. Tricco, E. Lillie, W. Zarin, K. K. O'Brien, H. Colquhoun, D. Levac, D. Moher, M. D. Peters, T. Horsley, L. Weeks *et al.*, "Prisma extension for scoping reviews (prisma-scr): checklist and explanation," *Annals of internal medicine*, vol. 169, no. 7, pp. 467–473, 2018.

[102] P. Van Damme, J. Fernandez-Breis, N. Benis, J. Minarro-Gimenez, N. de Keizer, and R. Cornet, "Performance assessment of ontology matching systems for fair data," *Journal of Biomedical Semantics*, vol. 13, 2022.

[103] M. Van Lent, W. Fisher, and M. Mancuso, "An explainable artificial intelligence system for small-unit tactical behavior," in *Proceedings of the national conference on artificial intelligence*. Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; 1999, 2004, pp. 900–907.

[104] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, and I. Polosukhin, "Attention is all you need," in *Advances in Neural Information Processing Systems*, vol. 30. Curran Associates, Inc., 2017.

[105] W3C, "W3C Semantic Web - Web Ontology Language (OWL)," https://www.w3.org/OWL, 2012, accessed: 2023-01-18.

[106] ——, "W3C Semantic Web - Resource Description Framework (RDF)," https://www.w3.org/2001/sw/wiki/RDF, 2014, accessed: 2023-01-18.

[107] ——, "Semantic web," https://www.w3.org/standards/semanticweb/, 2015, accessed: 2022-07-11.

[108] ——, "The JSON-LD Vocabulary," https://www.w3.org/ns/json-ld, 2020, accessed: 2023-01-18.

[109] ——, "World wide web consortium," https://www.w3.org/, 2021, accessed: 2022-07-11.

[110] W3C Community Group, "Building topology ontology," https://w3c-lbd-cg.github.io/bot/, 2021, accessed: 2022-07-12.

[111] B. Wang, J. Fedele, B. Pridgen, A. Williams, T. Rui, L. Barnett, C. Granade, R. Helfrich, B. Stephenson, D. Lesueur, T. Huffman, J. Wakefield, L. Hertzler, and B. Poplin, "Evidence-based maintenance: Part i: Measuring maintenance effectiveness with failure codes," *Journal of Clinical Engineering*, vol. 35, no. 3, pp. 132–144, 2010.

[112] ——, "Evidence-based maintenance: Part ii: Comparing maintenance strategies using failure codes," *Journal of Clinical Engineering*, vol. 35, no. 4, pp. 223–230, 2010.

[113] ——, "Evidence-based maintenance: Part iv: Comparison of scheduled inspection procedures," *Journal of Clinical Engineering*, vol. 38, no. 3, pp. 108–116, 2010.

[114] ——, "Evidence-based maintenance: Part iii: Enhancing patient safety using failure code analysis," *Journal of Clinical Engineering*, vol. 36, no. 2, pp. 72–84, 2011.

[115] B. Wang, E. Furst, T. Cohen, O. R. Keil, M. Ridgway, and R. Stiefel, "Medical equipment management strategies," *Biomedical Instrumentation & Technology*, vol. 40, no. 3, pp. 233–237, 2006.

[116] E. Wang, H. Kang, and Y. Gong, "Generating a health information technology event database from fda maude reports," *Studies in health technology and informatics*, vol. 264, pp. 883–887, 2019.

[117] L. Wang, G. Jiang, D. Li, and H. Liu, "Standardizing drug adverse event reporting data," *Studies in health technology and informatics*, vol. 192, p. 1101, 2013.

[118] M. Wilkinson, S.-A. Sansone, E. Schultes, P. Doorn, L. O. Bonino da Silva Santos, and M. Dumontier, "A design framework and exemplar metrics for fairness," *Scientific Data*, vol. 5, no. 1, 2018.

[119] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, P. Cistac, T. Rault, R. Louf, M. Funtowicz, J. Davison, S. Shleifer, P. von Platen, C. Ma, Y. Jernite, J. Plu, C. Xu, T. Le Scao, S. Gugger, M. Drame, Q. Lhoest, and A. Rush, "Transformers: State-of-the-art natural language processing," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*. Association for Computational Linguistics, 2020, pp. 38–45.

[120] World Health Organization, "Standardization of medical devices nomenclature: International classification, coding and nomenclature of medical devices: report by the director-general," World Health Organization, Tech. Rep., 2019.

[121] ——, "Standardization of medical devices nomenclature: International classification, coding and nomenclature of medical devices: report by the director-general," World Health Organization, Tech. Rep., 2022.

[122] ——, "Nomenclature of medical devices," https://www.who.int/teams/health-product-policy-and-standards/assistive-and-medical-technology/medical-devices/nomenclature, 2023, accessed: 2023-03-23.

[123] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao, and J. Zhu, "Explainable ai: A brief survey on history, research areas, approaches and challenges," in *Natural Language Processing and Chinese Computing*, J. Tang, M.-Y. Kan, D. Zhao, S. Li, and H. Zan, Eds. Springer International Publishing, 2019, pp. 563–574.

[124] R. Xu and Q. Wang, "Large-scale combining signals from both biomedical literature and the FDA Adverse Event Reporting System (FAERS) to improve post-marketing drug safety signal detection," *BMC bioinformatics*, vol. 15, p. 17, 2014.

[125] H. Yu, S. Nysak, N. Garg, E. Ong, X. Ye, X. Zhang, and Y. He, "ODAE: Ontology-based systematic representation and analysis of drug adverse events and its usage in study of adverse events given different patient age and disease conditions," *BMC Bioinformatics*, vol. 20, p. 199, 2019.

[126] D. Zhou, D. Zhong, and Y. He, "Biomedical relation extraction: From binary to complex," *Computational and Mathematical Methods in Medicine*, vol. 2014, 2014.

[127] R. Zhu, X. Tu, and J. X. Huang, "Utilizing BERT for biomedical and clinical text mining," in *Data Analytics in Biomedical Engineering and Healthcare*. Academic Press, 2021, pp. 73–103.