5. BÖLÜM / CHAPTER 5

MEDICAL ONTOLOGIES

MEDİKAL ONTOLOJİLER

Dilek YARGAN*, Aziz F. ZAMBAK**

*Research Associate, Laboratory for Computational Ontology, Middle East Technical University, Ankara, Turkiye
E-mail: dilek.yargan@metu.edu.tr

**Assoc. Prof. Dr., Department of Philosophy, Middle East Technical University, Ankara, Turkiye
E-mail: zambak@metu.edu.tr

DOI: 10.26650/B/ET07.2022.005

ABSTRACT

Medicine is a complex discipline that bears on manifold disciplines in the life sciences. Therefore, data sources for medical informatics are not limited to health and hospital records, and medical literature; these sources include gene studies, clinical reports, pharmacology studies, and other studies in the life sciences. For this reason, the primary issue of medical informatics is the integration of data, the volume of which is increasing day by day, and which is derived from a wide variety of sources and produced in various formats. Data integration requires a *lingua franca*, i.e., data standardization, in medical informatics. Moreover, it should be ensured that the data are machine-readable so that the conditions of machine inferencing, which enables scientific discoveries, are met in order to establish data-driven endeavors in medicine. Thus, with the support of various technologies, information retrieval and extraction, knowledge management, and knowledge production can be undertaken in medical and clinical studies. Ontologies used in information systems for decades promise to achieve these five goals of medical informatics: data standardization, data integration, information retrieval and extraction, knowledge management, and scientific knowledge production.

Medical ontologies are effective technologies that are widely used in many different applications in biomedical information and knowledge management systems. They are employed to represent biomedical knowledge with reference to reality in computable formats. In order to better understand medical ontologies, it is necessary to understand ontologies as philosophy, ontologies as science, and ontologies as technique. Furthermore, their contribution to the development of medical ontologies should be appreciated: An understanding of ontology as philosophy is necessary for the correct understanding of important categories in medicine and the correct classification of reality related to medicine. Medical ontologies as science, on the other hand, should be up-to-date and benefit from the wisdom of the philosophy of medicine. Enriched theoretically by these two types of ontologies, medical ontologies, which are ontologies as technique used in medical informatics, are obtained at last.

Keywords: medical ontology, applied ontology, data standardization, knowledge production, knowledge representation

ÖZ

Tıp, yaşam bilimlerindeki birçok bilim dalı ile ilişkide olan karmaşık bir disiplindir. Bu nedenle, tıp bilişiminin veri kaynaklarını sadece sağlık ve hastane kayıtları, tıp alanyazını ile sınırlı değildir; bu kaynaklar gen çalışmalarını, klinik deney raporlarını, farmakoloji çalışmalarını ve yaşam bilimlerindeki daha başka çalışmaları da içerir. Bu nedenle, tıp bilişiminin başat meselesi hacmi her geçen gün artan, çok çeşitli kaynaklardan gelen ve çeşitli biçimlerde üretilen verilerin bütünleştirilmesidir. Veri bütünleştirilmesi ise tıp bilişiminde bir lingua francaya, yani veri standardizasyonuna ihtiyaç duyar. Bunun yanında, çağımızın veri güdümlü bilim yapma biçimini tıp dünyasında kurmak için bilimsel keşifleri mümkün kılacak çıkarımların sağlanabilme koşulu olan verilerin makine okuyabilir olmaları sağlanmalıdır. Bu sayede, çeşitli teknolojilerin desteğiyle enformasyon erişimi ve çıkarımı, bilgi yönetimi ve bilimsel bilgi üretimi makine destekli tıp çalışmalarında görülebilir. Enformasyon sistemlerinde on yıllardır kullanılan ontolojiler tıp bilişiminin bu beş hedefini—veri standardizasyonu, veri bütünleştirme, enformasyon erişimi ve çıkarımı, bilgi yönetimi, bilimsel bilgi üretimi—sağlamayı vadetmektedirler. Medikal ontolojiler, medikal öneme sahip nesnelerin, niteliklerin, süreçlerin ve bunlar arasındaki ilişkilerin gerçekliğe referansı ile biyomedikal bilgiyi makinelerde berimsel yapılarda temsil ederek biyomedikal enformasyon ve bilgi yönetiminde çok sayıda farklı uygulamada yaygın kullanılan etkili teknolojilerdir. Medikal ontolojileri daha iyi anlamak için felsefe olarak ontolojiler, bilim olarak ontolojiler ve teknik olarak ontolojilerin neler olduğu bilinmeli ve medikal ontolojilerin gelişimlerindeki katkılarına dikkat edilmelidir: Felsefe olarak ontolojiler tıpta önemli kategorilerin doğru kavranması ve gerçekliğin doğru sınıflandırılması için bilinmelidir. Bilim olarak ontoloji şeklinde üretilen medikal ontolojiler ise güncelliklerini korumalı ve tıp felsefesinden beslenmelidir. Bu iki çeşit ontoloji ile kuramsal olarak zenginleştikten sonra tıp bilişiminde faydalanılacak teknik olarak ontoloji olan medikal ontolojiler elde edilir.

Anahtar Kelimeler: Medikal ontoloji, formel ontoloji, veri standardizasyonu, bilgi üretimi, bilgi temsili

1. Introduction

Medical informatics can be defined as a discipline that develops the necessary technologies for recording, storing, sharing, and processing health data. Health data, on the other hand, constitute our health records, such as vaccination history, allergy information, surgery history, and previous and current drug and treatment histories. These data are processed for the purposes of establishing decision mechanisms, such as diagnosis, treatment, disease prevention as well as of establishing a system that will provide better healthcare services, such as by setting clinical standards, perfecting the data sharing between health centers and professionals, and designing better treatment methods. Moreover, thanks to Big Data and cloud technology, we now have access to the resources and technologies needed to conduct precision medicine studies, which have been targeted for many years in medical informatics. Precision medicine is a medical informatics field that profiles patients for the prevention and treatment of diseases according to their disease-specific conditions, their medical histories, and their unique characteristics, such as genetics, environment, and lifestyle. Since precision medical practices are part of medical informatics, it is not possible to consider medical studies to be limited to patients' medical reports. Medicine is essentially a highly sophisticated research field in the life sciences, and many fields in the life sciences interact with medical studies. For this reason, data used in medical research are not only obtained from electronic medical records or electronic health records, but also from gene studies, clinical trial reports, pharmacology studies, and other studies in the life sciences. At this point, we face a significant challenge in terms of information management: the diversity and complexity of the data. Moreover, among other data sources, even electronic health records include a wide range of data types, such as the patient's medical history, vaccine records, patient's statistical data (such as age and gender), and blood test results. Although these records are maintained according to the standards of the institution where they are kept, when the larger picture is considered, electronic health records may differ across health institutions. Therefore, it is the main goal of medical informatics to be able to benefit from all possible information sources at the same time.

Digitization of data is seen in all areas of the life sciences, where it is used, first and foremost, to deal with the volume, complexity, and diversity of data. Data digitization is also vital in providing data integration, ensuring machine-readable data for machine inferencing that facilitates scientific discovery, and establishing standardization for data representation that enables analysis of the results of machine queries, all of which are at the core of bioinformatics. Especially in the post-genomic era, ontologies have become one of the technologies capable of dealing with the abovementioned challenges of bioinformatics. Such systems that represent entities of a particular domain and their interrelations in a machine-readable format are favored for the following reasons. Firstly, ontologies create a standard language that allows domain experts to speak the same language. Secondly, they provide access to implicit knowledge by querying the domain information represented on machines. Therefore, it is crucial to develop ontologies in order for data in the life sciences to be usable by all stakeholders (Baclawski & Niu, 2006, p. 2). Accordingly, together with considering the data sources of precision medical studies, it is necessary to consider the ontologies in bioinformatics as including medical ontologies as well. Therefore, the term "biomedical ontologies" is used to refer to ontologies in the field of bioinformatics, which also include medical ontologies.[1]

Biomedical ontologies are the most studied and developed of all ontology endeavors (Yoo & Phillips, 2009, p. 219). Therefore, they have a distinct value in terms of guiding other disciplines beyond their exclusive study in the biomedical field. In the second part of this work, in order to better understand this indispensable field, we will examine the types of ontologies. In the third part, we will investigate the properties of biomedical ontologies and the purposes for which they are developed. In the fourth part, we will present examples from medical ontologies.

2. What is Ontology?

Defining ontology, which represents an adventure of two thousand five hundred years that has its foundations in philosophy and continues to exist as a technique in information systems today, requires examining it according to specified criteria. The criteria in this section are the purposes and methods of creating ontologies. However, before proceeding, it is necessary to talk about the umbrella term *ontology*.

Ontologies are, in the simplest sense, systems that determine and model representations of beings and relations between beings in a domain. These systems consist of two primary components: beings and relations. Since the names given to beings and relations can have different meanings, building an ontology begins with creating a lexicon, a list of common terms of a domain, so that a model is possible. A linguistic partnership is thereby provided between those who will use ontologies for either theoretical or practical purposes. Indeed, the identification of common terms is more than a list of essential words in a medical article: the list obtained will be used by all scientists in the domain for the same semantic purposes. After determining the beings (enzyme, anatomy, Golgi apparatus, red) to be used in mode-

1 When referring to ontologies that include medical ontologies, the term "biomedical ontologies" will be used; when referring to ontologies that are created to use only in the field of medicine, the term "medical ontologies" will be used.

ling, they are classified according to categories, such as substance, attribute, and process. The most important classifications in an ontology are taxonomies, in which beings are arranged according to subsumption or the part-whole relation.[2] Relations describe interactions between two or more beings in taxonomy.[3] Taxonomies take the form of a tree structure, but after other relations are defined, the structure transforms into a directed graph. Hierarchical structures in an ontology—taxonomies established according to the subsumption relation and meronomies established according to the part-whole relation—may differ according to the purposes of ontologies (Gibson & Stevens, 2009, p. 6); that is to say, there is no unique ontology of a domain. Designers can model the same domain in different ways according to their aims in establishing the ontology, their philosophical attitudes, and the application areas of ontologies.

Having sketched out what an ontology is, we will now examine ontologies according to the purposes and methods of their creation in three main parts: ontologies as philosophy, ontologies as science, and ontologies as technique. Ontology as philosophy is the discipline that investigates being *qua* being and determines the limits of stating that something *is*; ontology as science examines beings and their interrelations in light of scientific findings; finally, an ontology as technique is a technology that is created for knowledge representation in machines. In order to better understand medical ontologies, it is necessary to master these three distinctions.

2.1. Ontology as Philosophy

Ontology is a philosophical discipline that studies being *qua* being. This discipline seeks answers to questions such as what are to be called as "beings", what kinds of beings exist, what categories of beings there are, what the ontological status of numbers would be, whether or not attributes would be being independent of substance, and what the modes of existence would be. These inquiries have been on the agenda of philosophers for two thousand five hundred years, and according to the theories of the philosophers, the answers to these questions vary greatly. Some philosophers investigate what categories of beings exist within the boundaries or within the freedom of natural language. This ontological endeavor is called traditional ontology or pure ontology (Sadegh-Zadeh, 2012, p. 688). On the other hand, some philosophers investigate modalities and features of being by doing abstractions at different levels within the boundaries of formal systems and consider existence within an axiomatic system. This ontological endeavor is called formal ontology. As we can see from these examp-

2 For example, in the expression "Humans are animals," the class "Animal" includes the class "Human."

3 For instance, the relation "to participate" in a cell ontology relates the beings "cell" and "apoptotic process" such that "A cell participates in an apoptotic process."

les, the ways of examining ontologies as philosophy differ in terms of questions and answers. For example, philosophical ontologies claim that the ontological status of attributes is secondary and depends on an independent substance, whereas others argue that these beings exist on their own. Philosophical formal ontologies focus on how to express attributes in formal languages. For example, a formal ontology theory that accepts attributes as universals can express the axiom of universals as $\forall x\ (Universal(x) \leftrightarrow \exists y\ exemplifies(y, x))$.

2.2. Ontology as Science

Ontology as science is a scientific study that investigates beings and their interrelations in a particular domain. Based on this definition, it is commonsensical to ask the following question by looking at a certain health-related ontological structure: Is there a disease?. However, at this point, a philosopher asks the question: What is the definition of disease?. It is possible to give dozens of definitions of disease: dysfunction or loss of function, the disorder of a structure, infection, inflammation, and others. However, these definitions may be valid for certain diseases, but false for others. As a science, ontology is concerned with what and how we can classify a particular domain. Therefore, the questions of whether a disease exists, whether it is called a disease, and whether it is a spatial entity are within the scope of philosophical ontologies. How to determine the categories, how to construct definitions, how to customize relations, how to employ facts such as part-whole, time and boundary and related guidance come from philosophical ontologies. On the other hand, if a disease ontology is to be established, the question of the categories into which schizophrenia and COVID-19 fall is within the scope of ontologies as science. Accordingly, when we want to create an ontology as science, we need to determine the concepts from the most basic to the most up-to-date in the literature, while we do not need to concern ourselves with their counterparts in reality or whether they *really* exist. Therefore, by accepting the existence of pathogenetic mechanisms, genes, and regulatory processes as *a priori*, we begin to form medical ontologies (Sadegh-Zadeh, 2012, p. 736).

There are many purposes in establishing ontologies as science, but three important reasons are sufficient for this chapter. First of all, the formation of the terminology of a scientific domain is essential. Creating a shared vocabulary prevents naming the same being with different names and calling different beings by the same name. A lexicon, a thesaurus, and a glossary should be prepared in order to maintain a shared language in science. All these structures present a list of concepts that will be used in ontology building. The second reason is that taxonomies updated with scientific developments show updates to theories. When given a taxonomy, it is effortless to determine the lower and upper classes of the beings and their relations with other beings. For this reason, ontologies as science are part of making science. Finally, and this is the reason for the existence of this book, scientific ontologies are created to serve purposes such as automatic information extraction, control of the consistency of the created scientific system, and sharing and integrating scientific data by providing knowledge representation of a domain in machines. Scientific ontologies created to serve medical studies pass through certain stages and are used in medical information systems.

2.3. Ontology as Technique

The increase in data volume, diversity, and complexity has led information systems to certain dead ends in terms of processing and sharing data. The differences in knowledge representation cause petabytes of non-processable data. For example, biology, which is now data-driven, does not suffer from a lack of data production or access to data, but has difficulties in data integration and analysis (Kelso, Hoehndorf & Prüfer, 2010, p. 347). However, it is desirable for researches to be empowered to make meaningful discoveries in this field by processing data from different sources and in various formats. Among the main objectives of information technology is representing knowledge in a way that machines can interpret. Ontologies as technique are tools that model a domain so that machines can process the structured data (Gaudet, 2018, p. 1). The most important feature that distinguishes ontologies from other representation systems is that they are software artifacts that serve to represent knowledge by creating a shared language in machines. These ontologies are called *applied ontologies*.

It is possible to examine ontologies as technique roughly in two classes: domain ontologies and domain-independent ontologies. Domain ontologies are systems that represent entities[4] and relations between entities in a particular domain in machines. They are divided into two types according to the purpose of their establishment. *Application ontologies* are ontologies created for specialized information management purposes such as data storage, sharing, information extraction, and information representation. *Reference ontologies*, on the other hand, are ontologies that contain the knowledge representation of the most up-to-date studies in a specific domain. They can be used to create application ontologies or for application in information systems. Created for pragmatic reasons, application ontologies can use more than one reference ontology, contain additional axioms that are not included in these ontologies, and in some cases, data from certain databases in order to serve the purpose of their establishment (Gibson & Stevens, 2009, p. 16). The second class, domain-independent ontologies, also called *upper-level ontologies*, aim to present the categories of all entities and

4 In order to distinguish it from the philosophical use of the term "being," we prefer "entity" to refer to things in a domain in ontologies as technique.

the relations between them. The need for domain-independent ontologies is to establish semantic interoperability among reference and application ontologies by introducing categories that are common across all domains. For example, crossing over is a process; the knowledge that this entity is a process must come from a study that examines entities with more general categories. The basics of the categories of being, such as space-time, part-whole, and border, are derived from philosophical discussions. Hence, philosophical theories take place in the construction of domain-independent ontologies. On the other hand, the use of these theories in building an ontology is also shaped by the requirement that ontologies must be consistent with scientific facts and that these ontologies are created to be used in information systems. What we mean here is that since our goal is to build a framework ontology that can be used in all areas, strict philosophical approaches, such as idealism or nominalism, should not be adopted. Finally, upper-level ontologies can be used in many areas—in medicine or education—as they serve as a framework for reference and application ontologies, thereby enabling the sharing of knowledge bases.[5]

There are several stages in the construction of ontologies as technique. The first of these is to determine the entities of the domain for which the ontology is supposed to be built. At this point, there are several sub-stages: identifying standard terms, creating a thesaurus, and preparing a glossary. Thus, the first step toward structuring all entities has been taken. Next, a taxonomy of entities and a list of relations are created. Taxonomies provide a *structural* glossary for data analysis and computation of a particular domain (Sadegh-Zadeh, 2012, p. 697). The taxonomy itself is not an ontology: in ontologies, the correct model of reality is reached by introducing relations between entities, relations that are different from subsumption or part-whole, namely by introducing the computable semantic structure into the taxonomy (Bodenreider & Stevens, 2006, p. 257; Chute, 2005, p. 173). For example, the enzyme is a protein (the "Protein" class subsumes the "Enzyme" class), and a reaction type is an entity (the class of all entities subsumes the reaction type). There is a "has a function" relation between the enzyme and the reaction type: an enzyme has the function of catalyzing a particular reaction type. All the stages passed through so far—design and development—are the same stages that one has to go through when building an ontology as science. The next steps, namely the coding of this representation in such a way that machines can read it, fall within the realms of building an ontology as technique (Stanescu et al. 2011, p. 47). In order to be represented on the machine, definitions that are in natural language must be represented in a formal language. Just as in philosophical formal ontologies, the entire system is to be expres-

5 It is also possible to classify ontologies in various ways (see Stanescu et al., 2011, p. 47; Bürger, Simperl & Tempich, 2014, p. 189; Gibson & Stevens, 2009, p. 13; Guzzi, 2018, p. 809.)

sed in a formal language. These expressions are called axioms, which define the restrictions of entities and clearly state some implicit facts (Guzzi, 2018, p. 809). The fact that "all instances of carboxylic acid are also instances of carbon oxoacid" can be deduced from a taxonomy of a chemistry ontology and can be formalized as $\forall x\ (CarboxylicAcid(x) \rightarrow CarbonOxyacid(x))$; From the definition of other relations, the fact that hydrogen and carbon are components of hydrocarbons can be deduced and formalized as $\forall x \forall y\ (Hydrocarbon(x) \land hasPart(x, y) \leftrightarrow Hydrogen(y) \lor Carbon(y))$ (Hastings, 2017, p. 7). Therefore, when an ontology as science, say a medical ontology, is represented on machines with axiomatic theories of medical entities and their interrelations, this ontology becomes an ontology as technique. Many formal languages from first-order logic to descriptive logics can be used to represent ontologies on machines. After the axioms are determined and translated into a formal language, an ontology editor, such as Protégé, is used to edit the entire structure on the machine so that the whole structure becomes machine-readable. Then, such an ontology is ready to be used for purposes such as knowledge representation, measuring the consistency of knowledge representation, sharing knowledge bases, automatic information extraction, or presenting a formal framework (Stanescu et al., 2011, p. 45).

3. Medical Ontologies

It can be observed that humans have been in a struggle to categorize beings since the beginning of their intellectual efforts. The question that started philosophy of what *arche* is, namely what the essence of everything is, investigates the nature of life. For this reason, it is very appropriate to say that the first classifications in history are about life. As such, the life sciences are richer in classification than any other science (cf. Chute, 2005, p. 167).

Until the 1990s, all classifications in medicine were stored in medical records, biological and biomedical data, inventions, research articles, and hospital archives (Gaudet, 2018, p. 1). Today, thanks to Big Data, most sciences, if not all, are data-driven. Therefore, all those classifications must somehow be accessible through machines, which requires that all the sources that store domain knowledge must be machine-readable. As the life sciences are the pioneers of the data-driven era, bioinformatics is an interdisciplinary field of research that aims to draw meaningful conclusions by analyzing complex data from various the life sciences sources (Blanco, 2014, p. 298). On the other hand, one of the biggest obstacles to analyze such data is that the data is not recorded/used in a shared language. Although, the classifications in the life sciences have been offering a shared language for centuries, since the 2000s, especially during genome studies, each laboratory created and used its own nomenclature, which has caused the problem of knowledge bases and/or databases not being shared among the researchers. To address this problem, ontologies have frequently been used in modeling data since the beginning of the post-genomic period (Gibson & Stevens, 2009, p. 3; Kelso, Hoehndorf & Prüfer, 2010, p. 369).

As mentioned before, medicine is a complex discipline that is in contact with many other disciplines. If medical informatics refers to all informatics research and applications that support medical research fields, then the data sources of medical ontologies should include electronic medical records, laboratory reports, visual data, drug prescriptions, research articles, and datasets from related biological sciences. Today, finding, producing, and storing data are no longer issues, the current issue is to integrate all the data coming from different sources in a meaningful way and make them analyzable. Medical ontologies, then, support the integration of data and information stored in many discrete databases; form a formal basis of interoperable biological and medical systems; ensure that data on patients and diseases are easily shared, reused, and processed by the relevant units, and ultimately, are used to generate knowledge by bringing together different systems and querying those large data pools (Stanescu et al., 2011, p. 51; cf. Shah & Musen, 2009, p. 445).

Before creating an ontology, it is necessary to determine why the ontology is to be produced, which entities should be used, who will use this ontology, and when and for how long the ontology will be used (Baclawski & Niu, 2006, pp. 280 – 281).[6] For the first stage, suppose that the ontology will be designed for supporting decision-making processes. Having this aim in mind, in the second stage, we need to determine the entities to be used. For this, we should start from the terminologies in the literature first. However, when we refer to the definitions of the terms used in the literature, we may encounter situations that we do not want: due to its complex nature, medicine is full of different commentaries and disagreements about which entities exist (Sadegh-Zadeh, 2012, p. 711).

Some of the questions discussed in the philosophy of medicine are these: Is medicine an art or applied science? Do diseases exist independently of us, or are they just discoveries? Are autoimmune responses pathological responses or hypothetical structures? Should we consider the accuracy or the benefit of medical knowledge? What is the nature of the human mind, and are there mental illnesses? (Sadegh-Zadeh, 2012, p. 687). Although they remain within the pursuit of philosophy, the answers to these questions have a critical role in building a medical ontology. So, under which category should include the questionable entity of nomophobia—fear of losing one's cell phone connection—as well as entities that are not questioned, such

6 For the purposes of this work, we focus only on the second question.

as diabetes, and how should the differences between the relations between dubious entities and *really* existing entities be presented? Consider the following question: Are there diseases? Since most of the medical literature is about diseases, it would be counterintuitive to negate the existence of diseases. On the other hand, some leading medical doctors think the opposite: Armand Trousseau rejects their existence; Thomas Szasz calls them a "myth;" and Richard Koch labels them as fiction (Sadegh-Zadeh, 2012, pp. 717, 718, 730). So, how can a language be formed from various interpretations of different nosological, etiological, and treatment theories? As we mentioned above, medical ontologies can reflect philosophical discussions; on the other hand, they are the tools created for sharing information in the literature. For this reason, most of the ontologies make use of the most widely held definitions of medical terms. For instance, if not all, most medical ontologies use the International Classification of Diseases (ICD), which explains which diseases are classified and how they are coded. For example, according to the ICD-11 database, in the etiological classification, Corona is an infectious agent, a virus from the *Coronavirinae* subfamily of the *Coronaviridae* family. "U07.1" is written in the medical report of a person who dies from COVID-19, a disease caused by a variant of this virus. There is, then, a standardized language for use in both scientific classification and special cases in medicine. However, there is no *network of relations* in this structure: ontologies are required to explain the interaction between entities, that is, the relations that carry the semantic load of the entire system. The crucial issue about relations is their properties. For instance, the relationship between diseases and symptoms is causal, and no disease is the cause of itself (Stanescu et al., 2011, p. 46). The causality relation is transitive in some cases and intransitive in others.[7] Consequently, it is necessary to differentiate among the types of causes to examine diseases and symptoms within the chain of causes. For example, consider a disease caused by avitaminosis, say, pellagra, or the abovementioned nomophobia. They have different symptoms and different disease-causing factors; for instance, the former is due to the lack of a physical substance, while the latter is mostly related to social interactions. Even if the categorization of these two entities is correct, the same relation may have different functions for these entities; therefore, the relation types must be specified as well. In a nutshell, classification is just a part of building an ontology; a model inhabiting a semantic web is the structure that allows implicit knowledge in medical informatics. For this reason, the accuracy of semantic representations is mandatory for valid knowledge representation. No matter if we accept nosological realism or skepticism or take diseases as abstract entities,

7 Consider these two examples. (1) Deep wounds cause blood loss; Blood loss causes death: Deep wounds cause death. (2) Alcohol causes dehydration; Dehydration causes death: Alcohol causes death.
In the first example, the causality is transitive: Deep wounds will cause death directly; on the other hand, in the second example, the causal relation is intransitive.

tropes, or substances, it should be borne in mind that it is crucial to obtain valid inferences from the correct representations.

Hitherto, we have mentioned the data sources of medical ontologies and some discussions on building an ontology. Below, we will examine the areas where medical ontologies are frequently applied.

3.1. Data Standardization

The first stage of building an ontology is the standardization of the data to be used in the knowledge representation: an ontology is, first of all, a controlled vocabulary of the domain, which is understood by all the experts in that scientific field. In other words, it presents the terminology of the field (Masseroli, 2018a, p. 813). Since the terms will coincide with different meanings in different disciplines, it must be clearly stated what precisely a term signifies within that discipline. For this reason, terminologies include a lexicon of the relevant domain; a thesaurus that includes annotations, synonyms, and antonyms; and a glossary in which the meanings of the terms are presented (Arroyo & Siorpaes, 2014, p. 145).

The workload and complexity of providing a controlled vocabulary that combines highly detailed and complex terms in both medical informatics and biological sciences are obvious (Chute, 2005, p. 175; Bodenreider & Burgun, 2005, p. 213). The systematic representation of medical terminology lies behind the widespread use of medical ontologies (Stanescu et al., 2011, p. 51; Stevens & Lord, 2009, p. 741; Shah & Musen, 2009, p. 445). That is to say, medical ontologies as a theory form a *lingua franca* in the medical world, even if they are not used for higher computational purposes, such as information extraction.

In the life sciences, it is observed that the same entities are used in different forms, and different entities are expressed in the same form; entities with the same form, namely synonyms, are not differentiated with regard to their meanings, and there are formal inconsistencies (Masseroli, 2018b, p. 823; Stevens & Lord, 2009, p. 738). In the literature, we encounter situations like these: instead of chloramphenicol acetyltransferase, its acronym CAT is used; Bacillus subtilis is abbreviated as B. Subtilis; sigma K is shown in different typographies such as sigma(K) or sigma-K; and “Behcet’s syndrome” or “Old Silk Route Disease” can be used instead of “Behçet’s disease” (Nédellec, Nazarenko & Bossy, 2009, pp. 672, 674; Gaudet, 2018, p. 1; MeSH on Demand, n.d.). Another example is that although the relationship in “A protein activates B protein” and “A protein activates X gene” is the same in the form, the two relationships are different from each other (Shah & Musen, 2009, p. 454). It is evident to life scientists that the examples cannot be limited to these. The number of such variations tends to be directly proportional to the number of laboratories and researchers. According to Masseroli, data standardization, which is a permanent problem of the life sciences, is overcome by using ontologies in theoretical and applied fields (2018a). Thus, the creation of a standard for terminologies will build a common language that provides a clearer understanding of the studies conducted in various areas of medicine, such as anatomy, physiology, diseases, diagnosis, and protocols.

3.2. Data Integration

Since the seventeenth century, efforts have been made to access controlled vocabularies to solve the problem of data integration in the medical field (Gibson and Stevens, 2009, p. 4). It is obvious that the only way to solve this problem, especially in the post-genomic period, is to structure the data. We have already mentioned that ontologies as technique do the work. The most common reason why ontologies are used for computational purposes is that they offer annotations to entities (Hastings, 2017, p. 9). Annotations, or descriptive definitions (Masseroli, 2018b, p. 826), are the formalization of a controlled vocabulary. In other words, after the assignment of a unique identifier to each entity in the terminology, the structural representation of other explanations of the entities in the controlled vocabulary, such as their synonyms, the identifier in another ontology, the private notes of ontology designers, are provided (Kelso, Hoehndorf & Prüfer, 2010, p. 365). Thus, annotations can be considered metadata.[8]

Controlled vocabularies and annotations based on ontologies make database research and data interoperability easy (Shah & Musen, 2009, p. 445). A successful ontology, representing the possible names of an entity in different databases and its annotations in different ontologies, can combine heterogeneous data. Suppose that for research, databases of different hospitals and a disease ontology are to be used. In a disease ontology, diseases—regardless of the data in the databases—are classified, and based on the interrelations among the entities, a semantic web is established. Hospital databases are, on the other hand, usually created by scientists modeling data that they have based on specific criteria. Therefore, the geographical locations of hospitals, scientists' ability to model data, data volume, and taxonomy used in representations make significant differences in databases (Kelso, Hoehndorf & Prüfer, 2010, p. 347). What the different entities used in hospital databases correspond to within the ontology is determined from the annotations or taxonomy of the ontology (Hastings, 2017, p. 9). Employing statistical methods, biomedical data analyses and semantic similarity[9] analyses

8 There is no single form of annotations; for instance, there are semantic annotations (Bürger, Simperl & Tempich, 2014, p. 186) or functional annotations (Blanco, 2014, p. 302).

9 For example, "Melatonin is a safe and an effective drug for insomnia that does not require a doctor's

are also carried out at this level thanks to ontologies. Nevertheless, ultimately, ontologies are used to identify inter-entity relations that are in the use of the most diverse databases (see Stevens & Lord, 2009, pp. 743 – 744).

Here, we would like to mention the two most common sources of information of medical ontologies: *knowledge-based information*, which includes general medical knowledge obtained from the life sciences literature, and *patient-specific information*, which includes electronic health records of individual patients (Hersh, 1996, pp. 21 – 22; Stanescu et al., 2011, p. 51). The interaction between these two sources is clear: analysis of the data in the latter source contributes to the literature, and the information in the former source enables the determination of the methods to be used in the diagnosis and treatment of the diseases of individual patients and the formation of the information included in the second source. Although the aims of the ontologies prepared with these two information sources are different from each other, the use of a reference or an upper-level ontology ensures the data integration of different biomedical ontologies.

3.3. Information Retrieval and Extraction

The most significant feature of the post-genomic period is the increase in data volume, diversity, and complexity. Therefore, as we have already mentioned, various methods have been developed to access biomedical information represented in various forms in different sources and to use this information in knowledge management. Biomedical ontologies enable us to access information in two ways thanks to their controlled vocabularies. The first is information retrieval, which enables us to access structured data, and the second is information extraction, which enables us to access semi-structured and unstructured data.

Biomedical ontologies are one of the most preferred methods in information systems since they provide a model for indexing databases, and they can provide interoperability between databases (Shah & Musen, 2009, p. 447). Indexing is the mapping between controlled vocabularies and data. In other words, thanks to controlled vocabularies, which are integrated with software, the data in the databases are labeled to create the infrastructure for information extraction (Stevens & Lord, 2009, p. 740). PubMed[10] offers the oldest and most advanced scientific knowledge-based information database in the biomedical field. The abstracts and/or citations of all documents, books, and references to biomedical literature in PubMed are

prescription." "For people who have difficulty in falling asleep, a sleep hormone supplement is recommended." Using the traditional vector analysis method, the similarity between these two expressions is measured as "very unlikely similar." On the other hand, providing relevant domain knowledge, ontologies used in the analysis raise the similarity measure to "almost the same" (Yoo & Phillips, 2009, p. 221).

10 https://pubmed.ncbi.nlm.nih.gov

indexed by MeSH[11] (Medical Subject Heading) terms. Thus, the relevant information is obtained from the database in this way: the keywords expressed in natural language by researchers are matched with the relevant MeSH terms in PubMed's search engine (Chute, 2005, p. 167; Chen et al., 2005, pp. 15 – 16). In clinical information systems, the indexing tool used for information retrieval is the ICD[12] (International Classification of Diseases); the current version is the 11th edition. In essence, when integrated with various software in medical informatics, the controlled vocabularies offered by ontologies play an essential role in information retrieval and extraction.

3.4. Knowledge Management

Information retrieval and extraction can only be applied to existing information. When it comes to accessing implicit knowledge, these methods remain impoverished: while these two processes require a machine-readable infrastructure, access to implicit knowledge requires a machine-*understandable* infrastructure. To create a machine-understandable system, first of all, it is necessary to determine the entities and relationships between entities, and then to present these entity-relation structures with axioms in a formal system. The resulting system is an ontology.

It is clear enough that the "implicit" contains neither data nor information: both are represented and can be accessed by information retrieval techniques. What is implicit is knowledge: we can extract knowledge in the biomedical domain by processing the rules of a formal system in which the knowledge is represented in the semantic network offered by the ontology so that we relate and develop our existing biomedical knowledge through knowledge discoveries (Yoo & Phillips, 2009, p. 221).

Ontologies created with descriptive logics offer a knowledge base. A knowledge base is a system formed by the merging of the terminological component that represents the domain information with axioms and an assertion component that instantiates the particulars of the domain by employing these axioms. Many techniques and methods, such as natural language processing tools, various artificial intelligence algorithms, Bayesian approaches and other statistical methods for text summarization, knowledge discovery, and clinical decision support systems have been used in biomedical research and services related to both knowledge-based information and clinical information (Chen et al., 2005). Ontologies with knowledge bases form an enormous infrastructure by combining these two sources of biomedical information.

11 https://www.nlm.nih.gov/mesh/meshhome.html

12 https://www.who.int/classifications/icd

The terminological component of the knowledge base resembles a knowledge-based information database since it contains the domain information expressed in axioms; patient-specific information corresponds to the assertion component since it contains data from individual patients, which are represented in accordance with the terminological component. In this way, queries can be addressed to the biomedical knowledge base as a whole, hypotheses can be tested, and research can be conducted on individual patients.

The queries to be addressed to the knowledge base can be considered as reasoning: each new query is crafted to reach a result from the set of propositions presented in the formal system via the emergence of unknown relations (Masseroli, 2018a, p. 813). The inferences are not explicit statements expressed in the knowledge base; if so, this would be information extraction. Inferences are implicit knowledge, which is discovery. For example, from a query to our knowledge base, we can find out in which anatomical contexts, certain tissues and cells are located (Hastings, 2017, p. 10).

Clinical information includes many data sets from electronic patient records to personnel management documents. There are various systems, the intended uses of which differ, which make use of these data sets. Here are a few examples of such systems: patient record management systems that enable healthcare professionals to access the data they are looking for in a clinical data deluge directly and quickly (Chen et al., 2005, p. 15); hospital information system, including management, operation, and patient information of an entire hospital (Montani & Bellazzi, 2002, p. 80); clinical decision support systems that assist healthcare professionals in the diagnosis and treatment of patients. The ultimate goal of clinical information systems is to test and/or discover a medical relationship or feature and to examine its validity in the biomedical domain (Colombo, Merico & Gündel, 2014, p. 246). Among all other systems that integrate clinical information, clinical decision support systems are the information technologies that best support this ultimate goal.

Clinical decision support systems are possible by combining two important aims: (i) to standardize the immense volumes of data at the highest level where methods such as data mining, statistical analysis, and logic structures are to be applied and (ii) to provide healthcare professionals with a product that is “meaningful, intuitive and overlaps with their mental models and business organizations” (Colombo, Merico & Gündel, 2014, p. 245). For decision support systems that will serve a biomedical domain, logical expressions of the rules are required first; then with these statements, a standard of medical logic modules and decision support rules is established by using knowledge-based information in the literature (Chute, 2005, p. 168). In the next stage, designers adapt these logic modules to a specific domain.

Although designers use standardized terminology and logical structures, they build a decision support system using their own classifications. For this reason, healthcare professionals using the decision support system may not be able to use the correct statements to run the rules. In addition, attention should be paid to the differences between the terminology of the language used by healthcare professionals and the terminology of the literature. Since the terminology that dominates the literature itself is accepted as a standard, designers also use the structures in the literature in clinical decision support systems. The semantic difficulty we face derives from the fact that neither terminology nor taxonomy is used as a standard structure and neither contains enough diversity.

Basing a decision support system on ontology solves these problems. First of all, ontology entails the standardization of terminology and taxonomy. While a medical ontology standardizes relevant literature, it also includes various usage information; as such, it is expected to include the language used in practice. Secondly, the designers do not need to reinvent the wheel, as an ontology offers clear taxonomy. In addition to all these, there are already standard formal inference rules as ontologies present the knowledge representation in a formal language. Designers do not need to change these rules according to their own organization. Ontologies are a ready-made package to facilitate decision support systems. Thirdly, ontologies also mitigate the computational burden of decision support systems: they offer the opportunity to narrow the research space of enormous and complex knowledge bases of the biomedical domain as much as possible thanks to the constraints expressed by the axioms (Stevens & Lord, 2009, p. 741). Finally, and above all, the knowledge bases provided by ontologies allow queries about particulars (cf. Kumar & Smith, 2004). The assertions component, where knowledge bases hold formal information about particulars, can be thought of as records of individual patients or particular situations. When all the information of the individual patient is expressed in the relevant formal language, inconsistencies within the patient's information can be detected, and inquiries that scan the entire all knowledge base, including all the information about the patient, can be made. In sum, beyond being standardizers of terminology and taxonomy, ontologies should be understood as a technology that offers a medical logical structure upon which decision support systems can be built.

3.5. Scientific Knowledge Production

Lastly, we want to mention scientific knowledge production. Although the processes of knowledge production differ among sciences, the following guidelines are generally followed. (1) Scientists formulate a hypothesis, model, or set of models. (2) They gather all preliminary information about the hypothesis/model(s) from various sources and (3) measure the consistency of the hypothesis/model(s) against this background. (4) When necessary, they review and correct the hypothesis/model(s). (4.1) They choose the most consistent among the models. Thus, the hypothesis/model(s) that are consistent with the preliminary information is/are included in the information system and direct the subsequent studies (Racunas et al., 2004, p. 257; Shah & Musen, 2009, p. 448). As mentioned above, hypothesis/model consistencies can be determined via a knowledge base. However, this is more than a consistency-check process: determining the consistency of hypotheses, represented in a machine-understandable way in a knowledge base using information technologies, is today's way of doing science.

In order for a hypothesis/model to be represented in a machine-understandable way, first of all, the following must be specified: biomedical objects, processes, and their interrelations must be represented in a way that is machine-readable; the qualitative and quantitative limitations of the entities must be determined; the rules of relations must be specified, such as transitional or symmetrical (Racunas et al., 2004, p. 257). As expressions of constraints, rules are criteria for determining the consistency of the hypothesis/model. Therefore, along with these criteria, background information that will determine the consistency of the hypothesis/model must be machine-understandable. Thus, thanks to certain reasoners (which are mainly based on ontologies), the coherence and consistency of the hypothesis according to the background information are determined by the shared formalism and a program that understands this formalism.

It is evident that amid the complexity and multiplicity of biomedical data, scientists depend on machines in scientific knowledge production. In this broad and deep domain, to determine the hypotheses and the consistency of hypotheses, or to conduct "thought experiments," as Shah and Musen say (2009), the first thing to do is to implement data standardization that provides access to multiple information sources through shared vocabulary. The next step is to represent the data in a machine-understandable way so that the machines can process data from various information sources. However, the construction of formal structures that machines will use when processing data allows thought experiments to produce results on machines and the experimenter to understand the *reasons* for the possible results.

Recall the definition of ontologies as technique: ontologies are representations of a formal *axiomatic theory* (Shrager et al., 2007, p. 340) that defines entities belonging to a domain and connects them. The rules defined above coincide with the axioms in this theory, and the language of representation matches the logic system of this theory. Therefore, ontologies represent the most appropriate theoretical and computational structure for thought experiments (Shrager et al., 2007; Racunas et al., 2004) since they provide standardization of the data in the biomedical field with a holistic philosophical attitude in the first step in their development. At the same time, knowledge bases created to structuralize complex and extensive biomedical knowledge can be established with biomedical ontologies. Since all is transferred to the machine in a logic system, the biomedical knowledge base itself turns into a formal axiomatic theory. Eventually, this system provides the infrastructure for biomedical scientists to conduct thought experiments (see Racunas et al., 2004; Shah & Musen, 2009).

By the way, an addition is in order. In the nature of scientific studies, deductive and inductive approaches are intertwined. Ontological inferences are mostly deductive; for this reason, modeling uncertainty in information systems is also gaining in importance. For example, Bayesian approaches are often used to model uncertainty in biomedical studies and to make stochastic inferences (Baclawski & Niu, 2006, p. 364). However, the use of various statistical methods and machine learning algorithms does not mean that ontologies are excluded; on the contrary, a system that is a combination of a biomedical ontology and Bayesian networks draws stochastic implications alongside logical implications: in addition to queries addressed to the knowledge base, statistical decisions are also derived. Similarly, when we integrate ontologies into clustering algorithms, we can develop knowledge-discovery tools that are capable of presenting new perspectives on the domain represented (Popescu et al., 2009, p. 60), and text mining techniques work more efficiently thanks to the taxonomy and annotations provided by ontologies that present the context of biomedical texts (Yoo & Phillips, 2009, p. 246). Therefore, biomedical ontologies should primarily be prepared as a foundation, and both their computational and scientific competence should be increased by integrating them with other techniques.

4. Examples of Medical Ontologies

In the previous section, we mentioned the contributions of ontologies to medical informatics. In this section, we will include various ontologies used in medicine.

The GALEN (Generalised Architecture for Languages, Encyclopaedias, and Nomenclatures in medicine) Project, started in 1991 as a European Union project, was created to enable multilingual clinical information created and used by different groups to be shared and processed so that this information could be used in medical records, decision systems, and other clinical systems. For this purpose, the researchers aimed to create a dynamically expanding ontology using an automatic classification engine instead of static, monolithic terminologies (Rector, Rogers, Zanstra & van der Haring, 2003, p. 982). At the end of the project in 1999, the results of the studies were shared and presented to the world of medicine as *Open*GALEN,[13] a framework that could be further developed.

13 https://www.opengalen.org

GALEN's ontology is designed to be a model that is reusable, and domain- and language-independent so that the application areas of the ontology need not to be limited to applications in clinical systems. It is used for associating the terminology between different electronic health records; providing content for electronic health records and decision support systems; providing easy-to-use and fast access to decision support and electronic health records for the clinical user interface; assisting designers to write compatible decision-support health-records hypermedia systems from reusable components; organizing and indexing information effectively and flexibly for decision support, hypermedia and bibliographic resources for knowledge management systems; and, finally, making available structured information that meets the expectations of clinical users and accepting natural language input into structured knowledge systems (van der Haring, n.d.).

MENELAS[14] is a European Union project that aims to provide better access to all information related to patients' histories and to process this information more effectively by establishing a system design that analyzes medical texts in English, French, and Dutch (Bodenreider & Burgun, 2005, p. 223; Zweigenbaum, 1995, pp. 82 – 84). The concept and relation types in the MENELAS ontology have been determined by using terminologies, text analysis, interviews with physicians, and many other sources in the literature (Bodenreider & Burgun, 2005, p. 223). In the ontology, the aim is to isolate the variability and flexibility in a natural language and between the languages by providing isolation from the natural languages to which the sources belong.

The objectives of MENELAS are to provide standardization of the clinical data in the specified European languages and to develop a system that helps health professionals' daily practices by better processing the documents of patients represented in natural language (Menelas Project Top-Level Ontology, 2018). Accordingly, this ontology is used in natural language processing applications in medicine, specialized applications, according to the needs of end-users that require analysis from documents represented in the natural language, international clinical research, and text analysis (Zweigenbaum, 1995, p. 83).

The Foundational Model of Anatomy[15] (FMA), which was developed to increase the anatomical content of UMLS and later turned into a reference ontology, represents the phenotypic structure of the idealized human body (Stanescu et al., 2011, p. 53). This ontology identifies the structural relations between entities at the highest level of detail, including entities from the molecules that make up the body to those at macroscopic levels (Rosse & Mejino, 2003,

14 https://bioportal.bioontology.org/ontologies/TOP-MENELAS

15 https://bioportal.bioontology.org/ontologies/FMA

p. 479). Since anatomy, which is the basis of all biomedical areas, is associated with the inclusion of anatomical entities and relations by the FMA, it is inevitable to consider the FMA as a *foundational* tool to be used in all domains of medicine (see Rosse & Mejino, 2003, p. 480).

The FMA is designed to associate different anatomical representations in biomedical informatics, to reference a wide range of domain ontologies being prepared in bioinformatics, and to create a consistent representation system at all levels of anatomical and biological structures and organizations. This design is, therefore, a reference ontology for use in various fields, such as physiology, clinical medicine, developmental biology, and pathology, where anatomical knowledge is indispensable. It is also used to enrich the semantic networks of various medical ontologies by defining the concepts of anatomy and the semantic interpretation of spatial relations (Rosse & Mejino, 2003, p. 496).

The Human Disease Ontology,[16] or Disease Ontology for short, was created in 2003 to improve the data of the NUgene Project conducted by the Genetic Medical Center at Northwestern University. (Bernasconi & Masseroli, 2019, p. 838). It is currently being developed by the Institute of Genomics at the University of Maryland School of Medicine as an ontology that provides a consistent, reusable and sustainable representation of human diseases, phenotypes, and entities related to human diseases or, briefly, as an ontology that standardizes human diseases (Disease Ontology: n.d.).

The terms of Disease Ontology are mapped with dictionaries, ontologies, and medical coding systems that are most commonly used in the medical field, such as MeSH, the ICD, NCI Thesaurus, SNOMED, OMIM, and UMLS so that the terms of disease and medicine are semantically integrated, and the conflicts in the classification of diseases are eliminated; therefore, it is no surprise that the first areas of use of Disease Ontology were structural annotations, a controlled vocabulary of diseases, and identification of diseases (Schriml et al., 2018, p. D955; Bernasconi & Masseroli, 2019, p. 841). Secondly, it is used as a reference ontology by inference systems because it is also designed to facilitate the connection of genetic and clinical data and symptoms, extracting data from various sources and finding relations between disease-related genes, symptoms, and signs (Bernasconi and Masseroli, 2019, p. 838; Schriml et al., 2018, pp. D955 – D956). Other uses of this ontology are diagnostic evaluation, clinical or experimental data comparisons, and graphical and statistical applications.

5. Conclusion

In conclusion, ontologies are established for various purposes in philosophy, science, and information systems. In terms of information and knowledge management, they are used

16 https://disease-ontology.org

in many fields, from data analysis to knowledge extraction, by providing knowledge representation on machines. The medical field benefits from ontology technologies due to gains in processable data from complex and diverse data sources and types. The data sources of medical ontologies include electronic health records, electronic medical records, laboratory reports, visual data, drug prescriptions, research articles, and datasets from medicine and related biological sciences. Applications of medical ontologies include supporting the integration of data and information stored in many discrete databases; creating a formal basis for interoperable biological and medical systems; ensuring that patient data and those pertaining to diseases are easily shared, reused, and processed by the relevant units; and generating knowledge by bringing together different systems and making queries to the resulting large databases possible. Medical ontologies are, therefore, effective technologies that are widely used in many different applications in biomedical information and knowledge management, where they represent biomedical knowledge in computational structures in machines, with reference to the reality of objects, qualities, processes, and the relationships between them. Such referencing can only be achieved in light of philosophy and science. Firstly, we want to emphasize that philosophical ontologies are essential for correct classifications of reality in the construction of ontologies. However, it is observed that the categories, such as universal, abstract, and properties, attributed to existence determined by traditional ontologies, as well as modes of existence, such as mereology, mereotopology, and time, are either misunderstood, underrated, or both by medical ontology designers (Herre, 2010, p. 389). Furthermore, medical ontology construction should benefit from the philosophy of medicine in order to arrive at the essence of entities. Secondly, the guidance of science is required for medical ontologies not only to obtain medical knowledge that is essential for building an ontology, but also to approve and verify the ontologies. In order to reach controlled vocabularies, taxonomies, semantic networks, axioms, and the like, the medical community must agree on each step of ontology building. However, constructing such communication tools upon which the medical community would agree is a tall order (see Sadegh-Zadeh, 2012, p. 733). Moreover, although common taxonomies are used, there may be differences in the definitions of entities, which can also alter semantic structures. An entity the taxonomic structure of which is obvious for a vast majority, such as "gene," can be defined very differently in the leading international genomic databases (Schulze-Kremer, 2002, pp. 179 – 180). Therefore, approval of ontologies as scientific work is essential. Besides this, scientists, along with ontology designers, have a role in ontology verification, which means the authentication of ontology construction, that is, ensuring that the definitions and structures used in ontology construction are correctly represented in the machine and that the formal system makes correct inferences. Ontologies demand the highest levels of technological progress; this is to be expected as, after all, "*Making the impossible* [is] *very difficult*" (van der Haring, n.d.).

References / Kaynakça

Arroyo, S. & Siorpaes, K. (2014). Ontologies and ontology languages. S. Miguel-angel (Ed.), In *Handbook of Metadata, Semantics and Ontologies* (pp. 141-155). World Scientific.

Baclawski, K. & Niu, T. (2006). *Ontologies for bioinformatics.* MIT Press.

Bernasconi, A. & Masseroli M. (2019). Biological and Medical Ontologies: Disease Ontology (DO). S. Ranganathan, M. Gribskov, K. Nakai & C. Schonbach (Ed.), In *Encyclopedia of bioinformatics and computational biology: ABC of Bioinformatics* Vol I (pp. 838-847). Elsevier.

Blanco, E. (2014). Metadata and ontologies for bioinformatics. S. Miguel-angel (Ed.), In *Handbook of Metadata, Semantics and Ontologies* (pp. 297-314). World Scientific.

Bodenreider O. (2004). The Unified Medical Language System (UMLS): integrating biomedical terminology. *Nucleic Acids Research*, 32 (Database issue): D267-270. doi: 10.1093/nar/gkh061. PubMed PMID: 14681409; PubMed Central PMCID: PMC308795.

Bodenreider, O. & Burgun, A. (2005). Biomedical ontologies. H. Chen, S. S. Fuller, C. Friedman & W. Hersh (Ed.), In *Medical informatics: Knowledge management and data mining in biomedicine* (pp. 211-236). Springer: Boston, MA.

Bodenreider, O. & Stevens, R. (2006). Bio-ontologies: current trends and future directions. *Briefings in bioinformatics*, *7*(3), 256-274.

Bürger, T., Simperl, E. & Tempich, C. (2014). Methodologies for the creation of semantic data. S. Miguel-angel (Ed.), In *Handbook of Metadata, Semantics and Ontologies* (pp. 185-215). World Scientific.

Chen, H., Fuller, S. S., Friedman, C. & Hersh, W. (2005). Knowledge management, data mining, and text mining in medical informatics. H. Chen, S. S. Fuller, C. Friedman & W. Hersh (Ed.), In *Medical informatics: Knowledge management and data mining in biomedicine* (pp. 3-33). Springer: Boston, MA.

Chute, C. G. (2005). Medical concept representation H. Chen, S. S. Fuller, C. Friedman & W. Hersh (Ed.), In *Medical informatics: Knowledge management and data mining in biomedicine* (pp. 163-182). Springer: Boston, MA.

Cimino, J. J. & Zhu, X. (2006). The practical impact of ontologies on biomedical informatics. *Methods of Information in Medicine*, 45(1), 124–135.

Colombo, G., Merico, D. & Gündel, M. (2014). Metadata and ontologies for health. S. Miguel-angel (Ed.), In *Handbook of Metadata, Semantics and Ontologies* (pp. 243-278). World Scientific.

Corcho, O., Fernández-López, M. & Gómez-Pérez, A. (2003). Methodologies, tools and languages for building ontologies. Where is their meeting point?. *Data & knowledge engineering*, *46*(1), 41-64.

Denaxas, S. C. (2017). Integrating bio-ontologies and controlled clinical terminologies: from base pairs to bedside phenotypes. C. Dessimoz & N. Škunca (Ed.), In *The Gene Ontology Handbook* (pp. 275-287). Humana Press: New York.

Disease Ontology, Retrieved May 5, 2020, from https://disease-ontology.org.

Fonseca, F. & Martin, J. (2007). Learning the differences between ontologies and conceptual schemas through ontology-driven information systems. *Journal of the Association for Information Systems*, *8*(2), 4. doi: 10.17705/1jais.00114.

Gaudet, P. (2018). The Gene Ontology. S. Ranganathan, K. Nakai & C. Schonbach (Ed.), In *Encyclopedia of bioinformatics and computational biology: ABC of Bioinformatics* Vol II (pp. 1-7). Elsevier. doi:10.1016/B978-0-12-809633-8.20500-1.

Gibson, A. & Stevens, R. (2009). Introduction to Ontologies. M. Popescu & D. Xu (Ed.), In *Data mining in biomedicine using ontologies* (pp. 1-21). Artech House.

Guzzi, P. H. (2018). Ontology in Bioinformatics. S. Ranganathan, K. Nakai & C. Schonbach (Ed.), In *Encyclopedia of bioinformatics and computational biology: ABC of Bioinformatics* Vol I (pp. 809-812). Elsevier. doi:10.1016/B978-0-12-809633-8.20490-1.

Haring, E. van der. *Open*GALEN Mission Statement. Retrieved June 1, 2020, from http://www.opengalen.org.

Hastings, J. (2017). Primer on ontologies. C. Dessimoz & N. Škunca (Ed.), In *The Gene Ontology Handbook* (pp. 3-13). Humana Press: New York.

Herre, H. (2010). The Ontology of Medical Terminological Systems: Towards the Next Generation of Medical Ontologies. R. Poli, M. Healy & A. Kameas (Ed.), In *Theory and applications of ontology: Computer applications* (pp. 373-391). Springer, Dordrecht.

Hersh, W. R. (1996). *Information retrieval: a health and biomedical perspective*. Springer Science & Business Media.

Hersh, W. R. (2005). Information retrieval and digital libraries. H. Chen, S. S. Fuller, C. Friedman & W. Hersh (Ed.), In *Medical informatics: Knowledge management and data mining in biomedicine* (pp. 238-275). Springer: Boston, MA.

Kelso, J., Hoehndorf, R. & Prüfer, K. (2010). Ontologies in biology. R. Poli, M. Healy & A. Kameas (Ed.), In *Theory and applications of ontology: Computer applications* (pp. 347-371). Springer, Dordrecht.

Kumar, A., & Smith, B. (2004). Enhancing GO for the sake of clinical bioinformatics. Retrieved from http://ontology.buffalo.edu/bio/ISMB/ISMB2004_Bio-ontology.pdf.

Masseroli, M. (2018a). Biological and medical ontologies: Introduction. S. Ranganathan, K. Nakai & C. Schonbach (Ed.), In *Encyclopedia of bioinformatics and computational biology: ABC of Bioinformatics* Vol I (pp. 813-822). Elsevier. doi:10.1016/B978-0-12-809633-8.20395-6.

Masseroli, M. (2018b). Biological and medical ontologies: GO and GOA. S. Ranganathan, K. Nakai & C. Schonbach (Ed.), In *Encyclopedia of bioinformatics and computational biology: ABC of Bioinformatics* Vol I (pp. 823-832). Elsevier. doi:10.1016/B978-0-12-809633-8.20491-3.

McCray, A. T. (2003). An upper-level ontology for the biomedical domain. *Comparative and Functional Genomics*, 4(1), 80-84.

Menelas Project Top-Level Ontology. NCBO BioPortal. (2018). Retrieved June 1, 2020, from https://bioportal.bioontology.org/ontologies/TOP-MENELAS.

MeSH on Demand, “Behcet Syndrome”, Retrieved April 13, 2020, from https://meshb.nlm.nih.gov/MeSHonDemand.

Montani, S. & Bellazzi, R. (2002). Supporting decisions in medical applications: The knowledge management perspective. *International journal of medical informatics*, *68*(1-3), 79-90.

Nédellec, C., Nazarenko, A. & Bossy, R. (2009). Information extraction. S. Staab & R. Studer (Ed.), In *Handbook on ontologies* (pp. 663-685). Springer: Berlin, Heidelberg.

Peleg, M., Yeh, I. & Altman, R. B. (2002). Modelling biological processes using workflow and petri net models. *Bioinformatics*, *18*(6), 825-837.

Popescu, M., Havens, T., Keller, J. & Bezdek, J. (2009). Clustering with Ontologies. M. Popescu & D. Xu (Ed.), In *Data mining in biomedicine using ontologies* (pp. 45-62). Artech House.

Racunas, S. A., Shah, N. H., Albert, I. & Fedoroff, N. V. (2004). HyBrow: a prototype system for computer-aided hypothesis evaluation. *Bioinformatics*, *20* (Suppl_1), i257-i264.

Rector, A. L., Rogers, J. E., Zanstra, P. E. & Haring, E. van der (2003). *Open*GALEN: Open source medical terminology and tools. In *AMIA Annual Symposium Proceedings* (p. 982). American Medical Informatics Association.

Rosse, C. & Mejino, J. L. V. (2003). A reference ontology for biomedical informatics: The Foundational Model of Anatomy. *Journal of Biomedical Informatics*, 36(6), 478-500.

Sadegh-Zadeh, K. (2012). *Handbook of analytic philosophy of medicine*, Philosophy and Medicine Series, Vol. 113, Springer: Dordrecht, Heidelberg, London, New York.

Schriml, L. M., Mitraka, E., Munro, J., Tauber, B., Schor, M., Nickle, L., … Greene, C. (2018). Human Disease Ontology 2018 update: Classification, content and workflow expansion. *Nucleic Acids Research*, 47(D1), D955–D962.

Schulze-Kremer, S. (2002). Ontologies for molecular biology and bioinformatics. *In silico biology*, *2*(3), 179-193.

Shah, N. & Musen, M. (2009). Ontologies for formal representation of biological systems. S. Staab & R. Studer (Ed.), In *Handbook on ontologies* (pp. 445-461). Springer: Berlin, Heidelberg.

Shrager, J., Waldinger, R., Stickel, M. & Massar, J. P. (2007). Deductive biocomputing. *PloS one*, *2*(4).

Stanescu, L., Burdescu, D. D., Brezovan, M. & Mihai, C. G. (2011). *Creating new medical ontologies for image annotation: a case study*. Springer Science & Business Media.

Stevens, R. & Lord, P. (2009). Application of ontologies in bioinformatics. S. Staab & R. Studer (Ed.), In *Handbook on ontologies* (pp. 735-756). Springer: Berlin, Heidelberg.

Tartir, S., Arpinar, I. B. & Sheth, A. P. (2010). Ontological evaluation and validation. R. Poli, M. Healy & A. Kameas (Ed.), In *Theory and applications of ontology: Computer applications* (pp. 115-130). Springer, Dordrecht.

Vrandečić, D. (2009). Ontology evaluation. S. Staab & R. Studer (Ed.), In *Handbook on ontologies* (pp.293-313). Springer: Berlin, Heidelberg.

Yargan, D. (2019). Formel Ontolojiler ve Betimleyici Mantıklar [Formal Ontologies and Descriptive Logics]. *Felsefe Arkivi- Archives of Philosophy, 51*: 271-281. https://doi.org/10.26650/arcp2019-5119.

Yargan, D. (2020). Üst Düzey Ontoloji İnşasındaki Felsefi Yaklaşımlar [Philosophical Aspects of Building Upper-Level Ontologies]. *Kilikya Journal of Philosophy,* (1), 32-50.

Yoo, I. & Phillips, W. (2009). Ontology Applications in Text Mining. M. Popescu & D. Xu (Ed.), In *Data mining in biomedicine using ontologies* (pp. 219-247). Artech House.

Zweigenbaum, P. M. (1995). Coding and information retrieval from natural language patient discharge summaries. M.J. Ladeira, M.F. Laires & H.I. Christensen (Ed.), In *Health in the New Communication Age* (pp. 82-89). IOS Press, Amsterdam.