Volume 2, Issue 9, September - 2024 **ISSN (E):** 2938-379X

COMPARATIVE ANALYSIS OF THESAURUS, WORDNET, ONTOLOGIES

Abdurakhmanova daughter of Nilufar Zainabuddin
Professor of the Department of Computer Linguistics and
Applied Linguistics of the National University of Uzbekistan
n.abdurakhmonova@nuu.uz

Kadyrova Zebo Gulboyevna
Teacher of the Department of Computer Linguistics and
Applied Linguistics of the National University of Uzbekistan
kadirovazebo5450@gmail.com

Allaberdiyeva daughter of Durdona Gurbanmurat
2nd Stage Graduate Student of the Department of Computer Linguistics and
Applied Linguistics of the National University of Uzbekistan
durdonallaberdiyeva39@gmail.com

Abstract
Today, it is an important issue to transfer linguistic models and speech capabilities of natural language to computer language and to solve language-related issues with the help of information technologies and methods. The stability of development of all languages in the information age, preservation of their national purity is one of the global problems. In the era of rapid development of computer technologies, interdisciplinary cooperation is important for the development of every science.

Introduction

Computer linguistics in Uzbekistan began to develop at the beginning of the 21st century with the efforts of A. Polatov and a number of important research works, including A. Rahimov, B. Mengliyev, S. Muhammedova, N. Jorayeva, N. It was carried out by Abdurakhmonova, Sh.Khamroyeva, A.Eshmo'minov, A.Akhmedova, G'.Abduvakhabov. As a result of the search for optimal methods of knowledge representation in the field of computer lexicography, a number of technologies based on semantics are being created. Such technologies include Ontology, Thesaurus, and WordNet.

Analysis of literature on the topic

Work on the creation and use of ontologies has made some progress abroad. Taking into account the active development of semantic technologies, in recent years, more attention has been paid to ontological modeling in the Russian scientific literature. In particular, among the works created in this field in Russia, the following can be included: Andreev A.M., Berezkin

Web of Teachers: Inderscience Research
webofjournals.com/index.php/1

D.V., Rymar V.S., Simakov K.V. 1, Gladun A.Ya., Rogushina Yu.V. [6], Dobrov B.V., Ivanov V.V., Lukashevich N.V., Solovev V.D. 2, Zagoruyko 3N.G. and others. At this stage, a number of broad ontologies containing several thousand concepts were created: OMEGA, SUMO, DOLGE, 4etc.

Research Methodology

Many literatures provide different definitions of the term ontology. Ontology as a term can be used in different ways. In philosophy and computer linguistics, the meaning of the term ontology is different.

The first meaning of ontology is a philosophical science that studies existence and the most general characteristics of existence ;

The second meaning of ontology is a computer resource representing the description of knowledge in a certain field 5.

The term ontology was used in the field of computers in 1967 by SHMealy 6. In addition, we can find different definitions of ontology used in the field of computational linguistics. We know that the most appropriate definition of the concept of ontology belongs to T. Gruber (1993). An ontology is a fully structured knowledge model that includes concepts, various relationships, rules, and axioms 7. Despite the differences in the various definitions of ontology, many authors agree on the main components of ontology.

The main components of the ontology are:

- classes or concepts;
- attributes;
- relationships;
- axioms;
- examples.

Web of Teachers: Inderscience Research
webofjournals.com/index.php/1

1Andreev A.M., Berezkin D.V., Rymar V.S., Simakov K.V. Ispolzovanie tehnologii Semantic Web v sisteme poiska nesootvetstviy v tekstax dokumentov.//URL: http://www.inteltec.ru/publish/articles/textan/rimar_RCDL2006.html

2Dobrov B.V., Ivanov V.V., Lukashevich N.V., Solovev V.D. Course track 16 presentation: "Ontology and thesaurus".

3Zagoruyko N.G. i dr. System "Ontogrid" for postroenia ontology //Computer linguistics and intellectual technology. Tr. a worker conference Dialog'2005. M., 2005. S. 146-152.

4 Nikonenko, A. A. (2009). Obzor baz znaniy ontologicalheskogo tipa.

5Andreev A.M., Berezkin D.V., Rymar V.S., Simakov K.V. Ispolzovanie tehnologii Semantic Web v sisteme poiska nesootvetstviy v tekstax dokumentov.//URL: http://www.inteltec.ru/publish/articles/textan/rimar_RCDL2006.html

6Gladun A.Ya., Rogushina Yu.V. Ontologies and corporate systems, Chast II//Korporativnye sistemy #1/2006 // URL : http :// www . management . com . ua / ims / ims 116. html

7Dobrov B.V., Ivanov V.V., Lukashevich N.V., Solovev V.D. Course track 16 presentation: "Ontology and thesaurus". //URL: http://download.yandex.ru/class/solovyev/plan.pdf ; (see also:

Any objects given in an ontology can be taken as *classes or concepts , and individuals* are objects belonging to ontology classes. Ontology units (classes and instances) can have properties - *attributes* . Each attribute usually has a name and a value and is used to store information specific to that entity. And *relations* represent the type of interaction between domain concepts. An example of a binary relationship is *a part-whole relationship.* The difference between a relationship and an attribute is that a relationship connects two classes, while an attribute describes the internal properties of objects through specific values. Among the relationships in ontology, the most important is *the taxonomic* relationship (class-subclass relationship, genus-species relationship, is-a). *Axioms are used* to write statements that are always true . They can be included in the ontology for various purposes, for example, they can be used to define complex constraints on attribute values, relationship arguments, to check the accuracy of the data described in the ontology, or to obtain new data.

Analysis and Results

In the design of ontologies, two directions can be distinguished, which have been developed separately for some time. The first concerns the representation of an ontology as a formal system based on mathematically precise axioms. The second direction was developed within the framework of computational linguistics and cognitive science. There, ontology is understood as a system of concepts that exists only in the human mind and can be expressed in a natural language (or other sign system) 8.

There are two alternative approaches to creating and exploring ontologies. The first is based on (formal) logic. The second (linguistic) is based on the study of natural language (in particular, semantics) and the construction of ontologies in large text arrays called corpora. Currently, these approaches are closely related to each other.

NVLukashevich distinguished three main principles of classification of ontologies:

• According to the level of formality
• According to content and composition
• According to the purpose of creation

According to the degree of formality, ontology is classified by NVLukashevich as follows. That is, (Figure 1) shows the spectrum of ontologies according to the degree of formality of the ontology and the use of certain formal elements. Each point corresponds to the presence of some basic structure in the ontology that distinguishes it from other points on the spectrum. A dotted line conditionally separates ontologies from other ontological resources

8Solovev V.D., Dobrov B.V., Ivanov V.V., Lukashevich N.V. Ontology i tesaurusy, Uchebnoe posobie. Kazan/Moscow. 2006. 157 c .

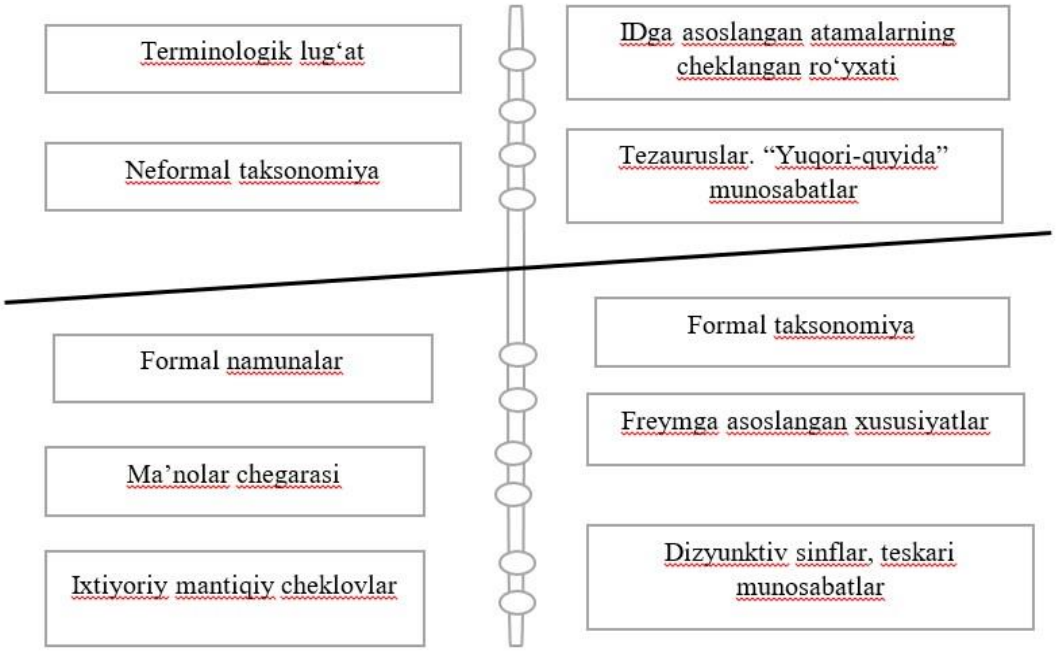

Figure 1. The ontology spectrum

The first point on the spectrum corresponds to controlled vocabulary, that is, a limited list of terms. Catalogs provide clear (not vague) interpretations of terms. For example, whenever we refer to the term "car", we use the same word regardless of what we are talking about in the context (which corresponds to some identifier in the dictionary): "washing machine", "car" , we mean "carriage machine", "garment machine" or other mechanical devices. Another ontology specification is *a dictionary* - this contains a list of terms with their meanings. Terms can have multiple meanings. Dictionaries are not sufficient for automatic processing by software agents. *A thesaurus is* a technology aimed at identifying semantic relationships between terms . Types of relationships specific to thesauruses: synonymy, hierarchical and associative relationships can be included 9.

The next point is " *formal taxonomies* ", these ontologies contain a precise definition of is-a (class-subclass) relationships. In such systems, the transitivity of the is-a relation is strictly observed: if class B is a subclass of class A, then every subclass of class B is also a subclass of class A. The class-element relation (is-InstanceOf) is satisfied by the following: if class B is a subclass of class A, then every instance of class B is also an instance of class A.

According to the purpose of creation, ontologies can be divided into the following 4 levels.

9Solovev V.D., Dobrov B.V., Ivanov V.V., Lukashevich N.V. Ontology i tesaurusy, Uchebnoe posobie. Kazan/Moscow. 2006. 157 c .

- representational ontologies,
- high-level ontologies,
- domain ontologies,
- are practical ontologies

The purpose of creating representational ontologies is to describe the domain of knowledge representation, to create a language for describing other lower-level ontologies. Example: Description of OWL language concepts using RDF/RDFS.

The goal of high-level ontologies is to create a single "correct ontology" that captures common knowledge for all disciplines and to reuse this ontology. There are several top-level ontology projects: SUMO, Sowa's Ontology, Cyc, and others 10.

The goal of a domain ontology is similar to that of higher-level ontologies, but the scope of interest is limited to a subject area (aviation, medicine, culture).

The purpose of practical ontologies is to describe a conceptual model of a specific task or program. They contain the most accurate information. Examples of practical ontological projects include TOVE, Plinius.

The classification of ontologies according to their content is as follows:

This classification is very similar to the previous one, but here the focus is on the actual content of the ontology. Ontologies and natural language text processing, automatic text processing, in particular, to solve the problems of information retrieval, it is necessary to match ontology concepts with a set of linguistic expressions (words and phrases). The procedure for comparing the ontology and the set of linguistic expressions can be done in different ways: First, the ontology can be preformed by logical classification, and then linguistic units can be assigned to its units. For example, Douglas Lenat, a well-known project leader in the field of CYC knowledge representation, 11 believes that the formalization and use of common sense knowledge is desirable for use in natural language text processing in particular. Verbalization of concepts assumes that the meanings of words can be confused. It is proposed to create a "top-down" ontology through logical analysis. At the same time, the names of the included concepts should (preferably) reflect the properties that form the basis of the division. Another problem with this approach is that when introducing linguistic expressions into a system of logically based concepts, the same word can correspond to many concepts depending on the context, which creates excessive polysemy of the lexical unit. Due to the unique characteristics of human language, even if a detailed domain ontology is constructed, ontologically linking knowledge in texts is a complex process.

According to the famous English linguist York Wills and other researchers, “All authors of ontology studies claim that concepts are the building blocks of any ontology, but we express concepts through words. In all known ontologies, words are used to represent concepts. Consequently, many non-verbal phenomena in the world cannot be modeled. We can describe

10Solovev V.D., Dobrov B.V., Ivanov V.V., Lukashevich N.V. Ontology i tesaurusy, Uchebnoe posobie. Kazan/Moscow. 2006. 157 c.

11Matuszek C., Cabral J, Witbrock M., DeOliveira J. An Introduction to the Syntax and Content of Cyc, http://www.cyc.com/doc/white_papers/AAAI06SS-SyntaxAndContentOfCyc.pdf

this phenomenon as Sapir-Whorf's ontological hypothesis, that is, what is not described in words cannot be reflected in ontology..." 12.

Hierarchical lexical resources such as WordNet also describe the lexicon of a natural language. Relationships between the meanings of words presented as separate units in a hierarchical network are synsets. The main feature of linguistic ontologies is that they are related to the meanings ("bound to semantics") of linguistic units (words, phrases, etc.). Linguistic ontologies cover most of the words in a language and at the same time have an ontological structure that is manifested in the relations between concepts. Therefore, linguistic ontologies can be considered a special type of lexical database and one of the types of ontology. Linguistic ontologies differ from formal ontologies by the degree of formalization.

The units of the medical field are also a specific hierarchy, which creates a hierarchy from top to bottom. From top to bottom, the hierarchy begins by finding the largest, main class, and its internal branches are identified. Medicine as a field is the basis of a large hierarchical group, that is, a subclass. As hierarchical units, all its internal fields are taken: *andrology, gastroenterology, gynecology, dietology, immunology, allergology, infectious diseases, cardiology,* etc. The above internal fields also form mutual internal groups.

Cardiology is a branch of medical science: it studies the structure, function and diseases of the cardiovascular system, their causes, mechanisms of development, specific course and diagnosis, and also develops methods of their detection, treatment and prevention 13.

As analyzed in our previous articles, there is a class of diseases in each branch of medicine. In the network of cardiology, the class of diseases is divided into two groups 14.

Figure 2. Class hierarchy

12 Solovev V.D., Dobrov B.V., Ivanov V.V., Lukashevich N.V. Ontology i tesaurusy, Uchebnoe posobie. Kazan/Moscow. 2006. 157 c.

13 Abdurakhmonova N. & Kadirova Z. (2023). In Computer Lexicography ontology. 150 p.

14 Tashkenbayeva EN Clinical cardiology: textbook - Samarkand: Medicine mirror, 2022. – 598 p.

Heart diseases;

Non-cardiac diseases.

The following names can be included in the class of heart diseases: ischemic heart diseases, cardiomyopathies, acquired and congenital heart defects, myocarditis, constrictive pericarditis, endocarditis due to infection, 15etc.

non-cardiac diseases includes respiratory system diseases accompanied by pulmonary hypertension, pulmonary artery thromboembolism, hypo and hyperthyroidism, diffuse connective tissue disease, anemia, hemochromatosis, amyloidosis, sarcoidosis, 16etc.

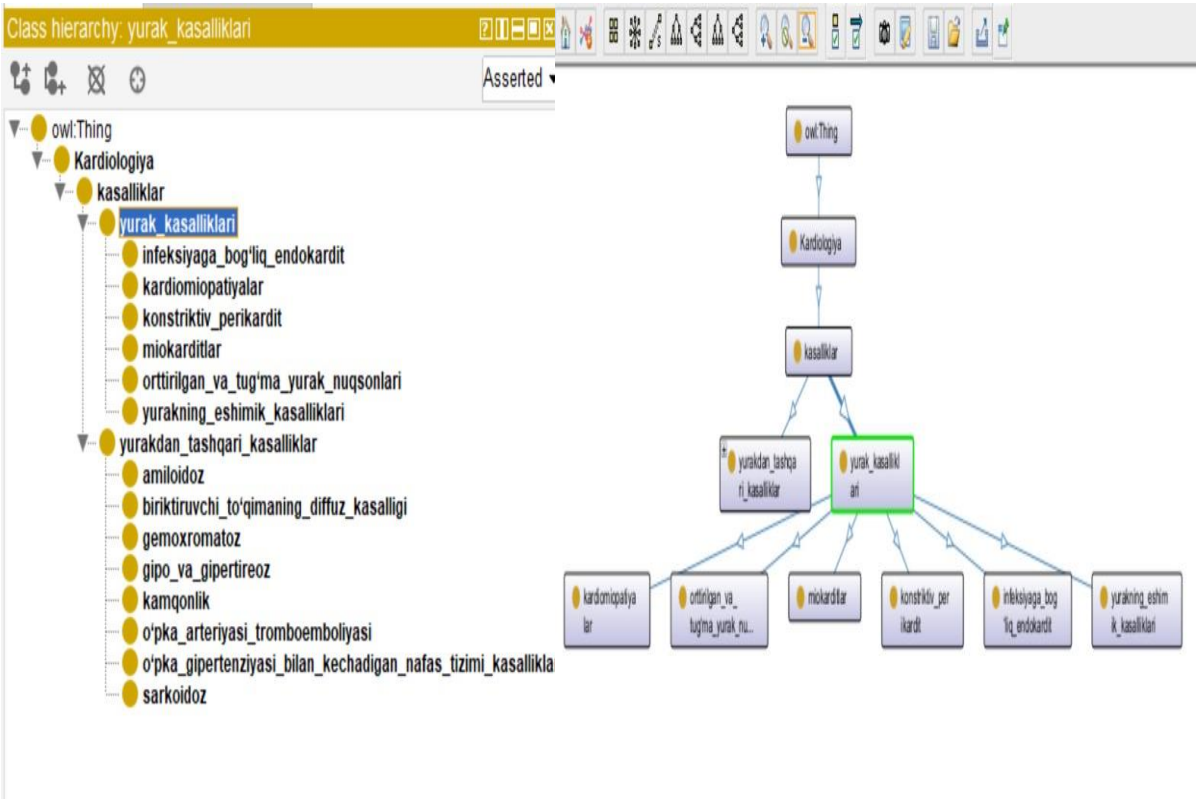

Figure 3-4. Representation of class hierarchy in OntoGraph

Ischemic heart disease is a common disease of the cardiovascular system ; It is accompanied by myocardial ischemia and coronary blood circulation disorders 17.

15 Tashkenbayeva EN Clinical cardiology: textbook - Samarkand: Medicine mirror, 2022. – 598 p.

16 Tashkenbayeva EN Clinical cardiology: textbook - Samarkand: Medicine mirror, 2022. – 598 p.

17 Sharipbay, A., Shirinova, R., Abdurakhmonova, N., & Kadirova, Z. (2023). ONTOLOGICAL KNOWLEDGE BASED MODELS REPRESENTING MEDICINE TERMINOLOGY. *Advances in Social Sciences and Economics* , *1* (1), 4.

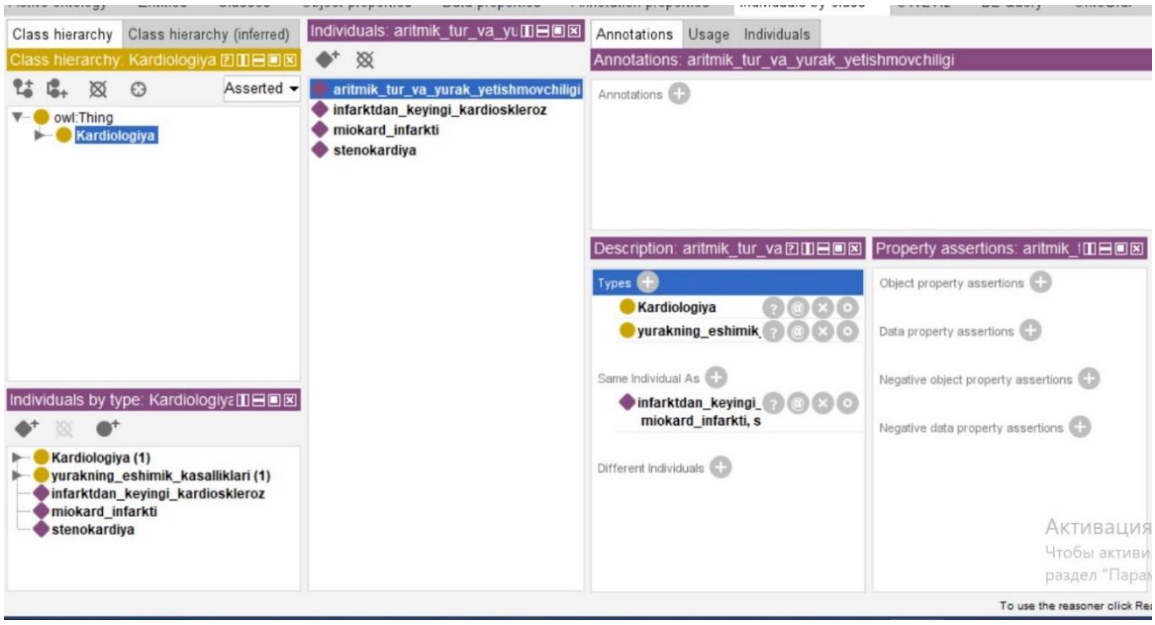

Figure 5. Representation of the elements of the class in the individual window

Ischemic heart diseases include angina pectoris, myocardial infarction, post-infarction cardiosclerosis, arrhythmic type, and heart failure (Fig. 5). individuals can be included.

Summary

Research on building ontologies and ontological models has been going on for a long time. All efforts are aimed at creating conditions for learning and further use of ontological engineering methodology.

Although WordNet, Thesaurus, and Ontologies are technologies based on the analysis of common semantic relations, they have different aspects. While thesauri represent concepts within a given topic and the semantic relations between them, WordNet is a resource belonging to the type of linguistic ontologies. And ontologies are a semantic technology based on defining concepts within a certain field and analyzing the relationships between them and determining the attributes of elements.

REFERENCES

1. Abdurakhmonova NZ Computational linguistics: textbook - Tashkent: Nodirabegim, 2021. - 394 p.
2. Matuszek C., Cabral J, Witbrock M., DeOliveira J. An Introduction to the Syntax and Content of Cyc, http://www.cyc.com/doc/white_papers/AAAI06SS-SyntaxAndContentOfCyc.pdf
3. Ruslan Mitkov, The Oxford Handbook of Computational Linguistics: Second Edition, 2022. DOI: 10.1093/oxfordhb/9780199573691.001.0001
4. Tashkenbayeva EN Clinical cardiology: textbook - Samarkand: Medicine mirror, 2022. – 598 p.

5. Andreev A.M., Berezkin D.V., Rymar V.S., Simakov K.V. Ispolzovanie tehnologii Semantic Web v sisteme poiska nesootvetstviy v tekstax dokumentov.//URL: http://www.inteltec.ru/publish/articles/textan/rimar_RCDL2006.html
6. Gladun A.Ya., Rogushina Yu.V. Ontologies and corporate systems, Chast II//Korporativnye sistemy #1/2006 // URL : http :// www . management . com . ua / ims / ims 116. html
7. Dobrov B.V., Ivanov V.V., Lukashevich N.V., Solovev V.D. Course track 16 presentation: "Ontology and thesaurus". //URL: http://download.yandex.ru/class/solovyev/plan.pdf ; (see also:
8. Zagoruyko N.G. i dr. System "Ontogrid" for postroenia ontology //Computer linguistics and intellectual technology. Tr. a worker conference Dialog'2005. M., 2005. S. 146-152.
9. Lukashevich N.V. Tesaurusy v zadachakh informatsionnogo poiska. – Moscow, 2010. – 388 c .
10. Nikonenko, A. A. (2009). Obzor baz znaniy ontologicalheskogo tipa.
11. Solovev V.D., Dobrov B.V., Ivanov V.V., Lukashevich N.V. Ontology i tesaurusy, Uchebnoe posobie. Kazan/Moscow. 2006. 157 c .
12. Abdurakhmonova N. & Kadirova Z. (2023). In Computer Lexicography ontology. 150 p.
13. Sharipbay, A., Shirinova, R., Abdurakhmonova, N., & Kadirova, Z. (2023). ONTOLOGICAL KNOWLEDGE BASED MODELS REPRESENTING MEDICINE TERMINOLOGY. Advances in Social Sciences and Economics, 1 (1), 4.
14. https://uz.wikipedia.org/wiki/Cardiology
15. https://uz.wikipedia.org/wiki/Heart_ischemic_disease.

Web of Teachers: Inderscience Research
webofjournals.com/index.php/1