Dual-Component Ontograph Visualization

A G Guraliuk[1], M L Rostoka[1], G S Cherevychnyi[2], D O Zakatnov[3] and T H Pavlysh[4]
[1]V O Sukhomlynskyi State Scientific and Pedagogical Library of Ukraine, Department of Scientific Information and Analytical Support of Education, 9/33 M. Berlynskoho st, Kyiv, 04060, Ukraine
[2]Taras Shevchenko National University of Kyiv, Department of Modern History of Ukraine, 64/13 Volodymyrska st, Kyiv, 01601, Ukraine
[3]V O Sukhomlynskyi State Scientific and Pedagogical Library of Ukraine, Department of Scientific Information and Analytical Support of Education, 9/33 M. Berlynskoho st, Kyiv, 04060, Ukraine
[4]Donetsk Law Institute of the Ministry of Internal Affairs of Ukraine, 21 Stepana Tilhy st, Kryvyi Rih, Dnipropetrovsk Region, 50065, Ukraine

ss_variant@bigmir.net, marilvross@gmail.com, gscherevichnyj@gmail.com, 23dazkum@ukr.net, tanyagrigorivna@ukr.net

Abstract. Today there are several generally accepted methods of storing data in an information network. One of the most modern is the presentation of data in the form of a semi-structured (with the possibility of fuzzy formalization) set – a knowledge base. Compared to databases, this view has several advantages: the ability to store complex heterogeneous information; the ability to expand and supplement the description of the subject area without reprogramming; visibility and accessibility of knowledge presentation to the user. But this advantage of them turns into a problem, which currently has been only partially solved. The data in the knowledge base can be in such a form that their presentation according to formalized rules and algorithms is impossible or ineffective. Currently, the most promising in terms of representing data structures is considered to be an ontological representation that copes well with displaying an arbitrary structure. Most of the visual representations of ontologies (Protégé in integration with OWLViz, OntoGraf, 3D Hyperbolyc, Tree, etc.) are images of nodes as a set of conditional points (small geometric shapes) connected by lines –ontographs. This approach to visualization is less informative and inconvenient to apply to the educational process.

The need to create new approaches to the presentation of information, implement its accessibility and efficiency of usage is becoming increasingly important. Along with traditional information support in the form of databases, knowledge bases have begun to develop at a tremendous pace, which, when used effectively, provide significant competitive advantages. An important feature of knowledge bases is the ability to work with approximate sets. The authors developed ontograph visualization technology, partially implemented on the Ontos.xyz web resource. This resource allows you to visualize the vertices of the ontograph, the connections between them and assign each set of contexts that are related to it. Each of the nodes of the ontograph contains a specific descriptive context. The main feature of the Ontos editor is the ability to assign each node a context of all types supported by the browser. Including html pages, web 2.0 resources, etc. Also, the proposed system implements the visualization of knowledge bases, consisting of two components: a navigator that determines the path to some node of the ontograph that has children and a visualization slider for this node that displays these elements. The knowledge base viewer is a slider that resembles a photo gallery slider with advanced navigation. The ontological approach to servicing knowledge bases can be not only a means of organizing knowledge. The knowledge bases developed on the basis of the ontological approach make it possible to work actively with knowledge, to solve problems associated with education, the development of artificial intelligence, decision-making systems, and many other areas where approximate sets can be used.

Keywords: Ontologies, Knowledge Bases, Tools for Formalizing Knowledge, Ontograph, Ontograph Visualization.

1. Research problem

In the conditions of the modern dynamic development of society, information be-comes an important strategic resource along with material and energy. The main thing that distinguishes the process of informatization is the acquisition of information on the status of one of the fundamental factors in the existence of mankind. However, there are certain problems with the definition of both the key concept of “information” and the essence of information in general.

“Information” is a fundamental element of information reality. Currently, there are several approaches to defining information. Among them, the most widespread are structural, functional-cybernetic, socio-oriented (anthropocentric), and recently – synthetic [1].

Each definition of information reflects an understanding of it from a different angle, for example, if we generally understand information, then this is a certain set of information about the world, events, a certain activity. It should also be noted that, although the form for presenting information is messages, not every message can be considered information. This is only the message that reduces the uncertainty in the area to which it relates. In addition, it is known that the concept of "information" includes the exchange of information in several aspects: between people; man and machine; automatic and automatic. From which it follows that the information should be formalized in a certain way.

The problems of formalizing information, implementing its accessibility and efficiency of use, are extremely acute. In our country, the public interest in them is fixed at the legislative level: so one of the priorities for the development of the information society in Ukraine is to ensure free access to information [2]. As a result, digital libraries have responded to the increasing information needs of users. The development of digital libraries means that people do not have to go to a building for some kinds of information, users still need help to locate the information they want. In a traditional library building, a user has access to a catalog that will help locate a book. In a digital library, a user has access to catalogs to find traditional library materials, but much of the information on, for example, the Internet cannot be found through one commonly accepted form of identification. This problem necessitates agreement on standard ways to identify pieces of electronic information (sometimes called meta-data) and the development of codes (such as HTML [Hypertext Markup Language] and SGML [Standard Generalized Markup Language]) that can be inserted into electronic texts.

An integrated individual service, primarily remote, is becoming the main activity of the library. The value of its information function is increasing. This, in turn, necessitates a systematic study of how modern man perceives, assimilates, uses information [3].

Moreover, the library should act as a kind of filter, selecting and providing users with scientifically reliable, verified, safe information. Ensure maximum pertinent of data received in response to user requests [4]. Unfortunately, the sociology of the reader, more broadly the consumer of information, is still underestimated by the library community. In many ways, it would help to draw up real ideas about what and in what form different categories of people from the library need, or what services they want from it. To solve the problems of information accumulation, databases are traditionally used. They simplify the solution of certain typical formalized tasks that are repeated under the conditions of operational functional activity. Databases contain, as a rule, developed regulations, models, and formalized algorithms. At the same time, the choice of the best option, for various reasons, cannot always be carried out according to formalized rules and algorithms. Therefore, along with traditional information support, knowledge bases have begun to develop at a tremendous pace, which, when used effectively, provide significant competitive advantages [5].

There are dozens, if not hundreds of definitions of the concept of a knowledge base. In the dictionary of modern computer vocabulary, the knowledge base refers to one or more specially organized files that store a systematic set of concepts, rules and facts related to a certain subject area. For example, knowledge base on the chemistry of hydrocarbons. The content of the knowledge base is drawn up, connected with each other, and presented in such a way that, based on it, it is possible to carry out reasoning and draw conclusions using special programs, obtaining information that may not be explicitly present in the knowledge base. To construct a knowledge base, artificial intelligence methods, special knowledge description languages and an intelligent interface are used [6]. Knowledge bases plays a critical role in many analysis tasks, both scientific and industrial, and is often the bottleneck to answering new and impactful macroscopic questions. In many scientific analyses, for example, one first needs to assemble a large, high-quality knowledge base of facts (typically from the literature) in order to understand macroscopic trends and patterns [7].

Currently, the most promising in terms of representing data structures is considered to be an ontological representation that copes well with displaying an arbitrary structure [8]. Let us explain this statement. Let there be some kind of database, for example, library readers. It contains database fields: number, surname, first name. There was a need for one of the students to assign the field "photo". To do this, you need to change the structure of the database table (or add a new table to the database). It is difficult enough for a non-programmer. If the organization of the data was presented in the form of a knowledge base, then it is enough to indicate the relationship between the node representing a particular student and the photo file, which is much simpler.

The aim of the research is to present a simple and intuitive knowledge base visualization technology for the end user, the use of which has a number of advantages over traditional display methods.

2. Approach

We propose to level the difference between a file and a directory and consider them as a certain set of objects of an arbitrary nature (any types of data that can be reproduced by computer technology), grouped according to some attribute. We consider such a collection as a node of a directed bipartite graph. Directional links are established for various objects (child and parent objects are defined). Thus, relationships arise between various objects of the form: "parent-child", "children of the same parent", "contextual content". The collection of all connections is a pyramidal network. To visualize network objects, we propose to combine the now classic data representation in the form of an ontograph which became classic with a slider to display the data of a certain node, which is a set of single-level data ("children of one parent"). In this case, the slider itself is equipped with hyperlinks for transitions to objects associated with the displayed objects (additional navigator).

This approach allows to combine the benefits of existing data display tools and render them with arbitrary granularity.

3. Analysis of Recent Research and Publications

According to K. Boskebeev, the only criterion of intelligence for information systems is the existence of mechanisms for working with knowledge [9], which are based on the theory of approximate sets. This theory is based on the assertion that with each object of the universe we associate some information (data, knowledge). Objects characterized by the same information are indistinguishable (similar) in terms of the information available about them. The indistinguishability relation generated in this way is the mathematical basis of the theory of approximate (rough) sets.

Approximate sets are based on the concept of an approximation space and the method of approximating a set and a family of sets. The approximation space is an ordered pair A = (U, R) in which U is the universe, and R is a kind of binary relation. The main advantages of using knowledge bases are: the ability to save complex heterogeneous information; the ability to expand and supplement the description of the subject area without reprogramming; visibility and accessibility of knowledge representation to the user. Such knowledge bases must have a fuzzy character of data ordering, close to human perception, and therefore intuitive to the user. Today, the ontological presentation of data is considered the most promising in this regard.

With the advent of the semantic network and semantic technologies, ontologies have emerged that have become one of the most modern paradigms for the representation of knowledge [10]. Ontology is a term borrowed from philosophy that refers to the science of describing the kinds of essences of various objects and the relationships between them. The use of computer ontologies was considered by T. Gruber, T. Jeffrey, Y. Ding, A. Evseev, V. Lapshin, V. Litvin, V. Lyubchenko, L. Naykhanov, A. Narignani, S. Nirenburg, S. Nirenburg, N. Noy, M. Popova, A. Smirnova, S. Subbotin, Y. Sure, D. Faure and many others.

4. Statement of Basic Material and the Substantiation of the Obtained Results

The approximation space presented in the universe under certain conditions differs little and can be described in terms of the ontological approach as a certain subject area. The knowledge base, as a tool that provides all three of the previously mentioned types of man-machine interaction, is part of the Universe U (Fig. 1.).

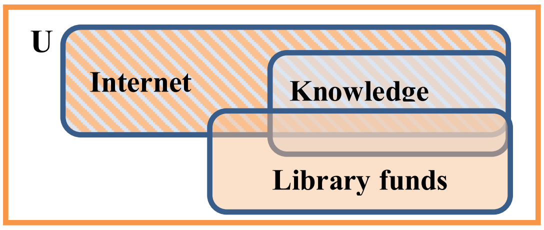

Fig. 1. Knowledge base of the organization in the "Internet library"

In modern realities, data on the Internet is an integral part of virtually any project related to the processing of digital information. The funds of the library (organization – the holder of information) can be both in digitized and in solid form. A knowledge base covering part of the information from the library repository should be able to obtain data from other sources (in general, it is the Internet, but options are possible). There are many mechanisms for creating knowledge bases, including those based on the ontological approach.

Their main task is to provide tools for formalizing knowledge and presenting them in the form of an accurate and comprehensive description. The main functionality of ontology editors is the creation of ontologies, their editing, replication, visualization. Many editors document ontologies, import and export them into various formats, and support ontology libraries. One of the definitions is deciphered by the term ontology, as a knowledge base that describes facts that are always assumed to be true within a particular society based on the generally accepted meaning of the dictionary used [11].

A formal description of an ontology in the form of a triplet is considered classical [4]: $O = <X, R, F>$ where O, in fact is ontology; $X = \{x1, x2, ... xn\}$ – a finite set of domain concepts; n is the number of elements in this set; $R = \{r1, r2, ... rm\}$ – a finite set of relations between the concepts of the subject area, m – the number of significant relations; F is a finite set of interpretation functions defined on concepts and ontology relations O. Visually, such a dependence can be represented in the form of an ontograph. By ontograph, in the sense of applying to the knowledge base, we mean a system X of formal elements (nodes of the ontograph), with given R connections (edges).

The authors developed ontograph visualization technology, partially implemented on the Ontos.xyz web resource. This resource allows to visualize the vertices of the ontograph, the relationships between them and assign each vertex sets of contexts that are interconnected with it. Context is the background, environment, setting, framework, or surroundings of events or occurrences. Simply, context means circumstances forming a background of an event, idea or statement, in such a way as to enable readers to understand the narrative or a literary piece. It is necessary in writing to provide information, new concepts, and words to develop thoughts.

Whenever writers use a quote or a fact from some source, it becomes necessary to provide their readers some information about the source, to give context to its use. This piece of information is called context. Context illuminates the meaning and relevance of the text, and may be something cultural, historical, social, or political.

Such an organization of the implementation of the structure of the vertices of the ontograph and the connections between them requires a description language. The most common language is OWL, a subset of which is the ontograph visualization language developed by us. But this will be discussed more detailed in a subsequent article.

For the final user, the top of the ontograph is associated with a separate file capable to contain information of several types. So when we click on the top (opening of a node) of the ontograph, we see various blocks that can contain heterogeneous information (graphic, video, text, sound, document output windows in MS Office, *.pdf, etc.). That is, we are considering such a set of data as a set of objects of arbitrary nature (any types of data that can be reproduced by computer technology), grouped according to a certain criterion. This combination is displayed as a node of a directed bipartite graph. Directional links are established for various vertices (nodes) (child and parent objects are determined).

While using such approach the need for the "classic" distinction between files and directories disappears. The same node can be perceived as a directory for child nodes and as a file for parent nodes. In this case, the number of parent nodes can exceed one, which removes the need to duplicate a file in different directories or create shortcuts.

Thus, relationships arise between various objects of the form: "parent-child", "children of the same parent", "contextual content". The combination of all connections is a pyramidal network. To visualize network objects, we propose to combine the data representation (which has become classic) in the form of an onto-graph with a slider to display the data of a specific node, which is a set of single-level data ("children of one parent"). In this case, the slider itself is equipped with hyper-links for transitions to objects associated with the displayed objects (additional navigator).

The proposed approach allows to combine the advantages of existing data display facilities and visualize them with arbitrary detail. A practical example of development implementation is given below. In the framework of the scientific research “Bibliographic and analytical support of the activities of the National Academy of Pedagogical Sciences of Ukraine on the scientific and methodological support of the modernization and reform of education”, conducted by the V.O. Sukhomlynskyi State Scientific and Pedagogical Library of Ukraine Sukhomlinsky was tasked with systematizing and publishing biographies of academicians of the National Academy of Pedagogical Sciences of Ukraine.

Figure 2 shows the ontograph of the president of the NAPS of Ukraine, V. Kremen. The balls indicate the nodes of the ontograph, and the connections are directed segments.

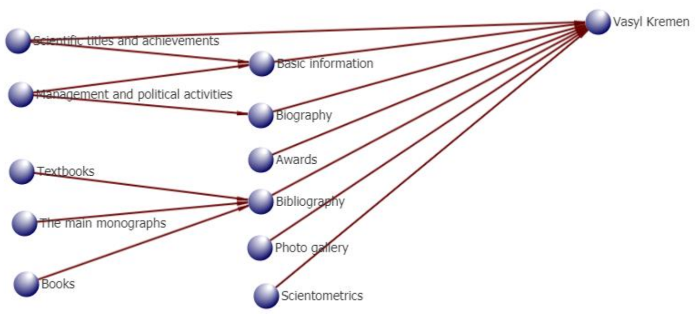

Fig. 1. Ontograph of the President of the NAPS of Ukraine V. Kremen

Each of the nodes of the ontograph contains a specific descriptive context. The main feature of the Ontos editor is the ability to assign each node a context of all types supported by the browser. Including html pages, web 2.0 resources, etc. (see **Error! Reference source not found.**).

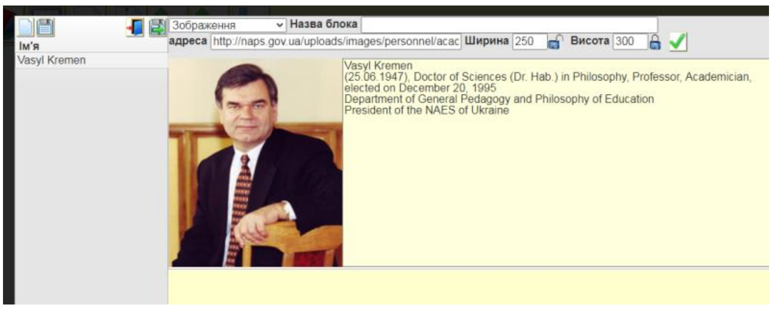

Fig. 2. The context of the node ”Biography“

Addition to the convenient ontograph editor, the proposed system implements a knowledge base visualization consisting of two components: a navigator that determines the path to an ontograph node that has children and a visualization slider for this node, which displays these elements.

When you click on an item to be displayed on the viewer’s screen, a pop-up window for the content of each item appears which may contain text, a picture or a hyperlink. A navigator can be an ontograph, a tree, navigation buttons, hyperlinks, etc. Its main goal is to determine the node whose visualization will be carried out. Another distinguishing feature of the development is the maximum approximation of the displayed node to a familiar electronic book (see Fig. 4, 5).

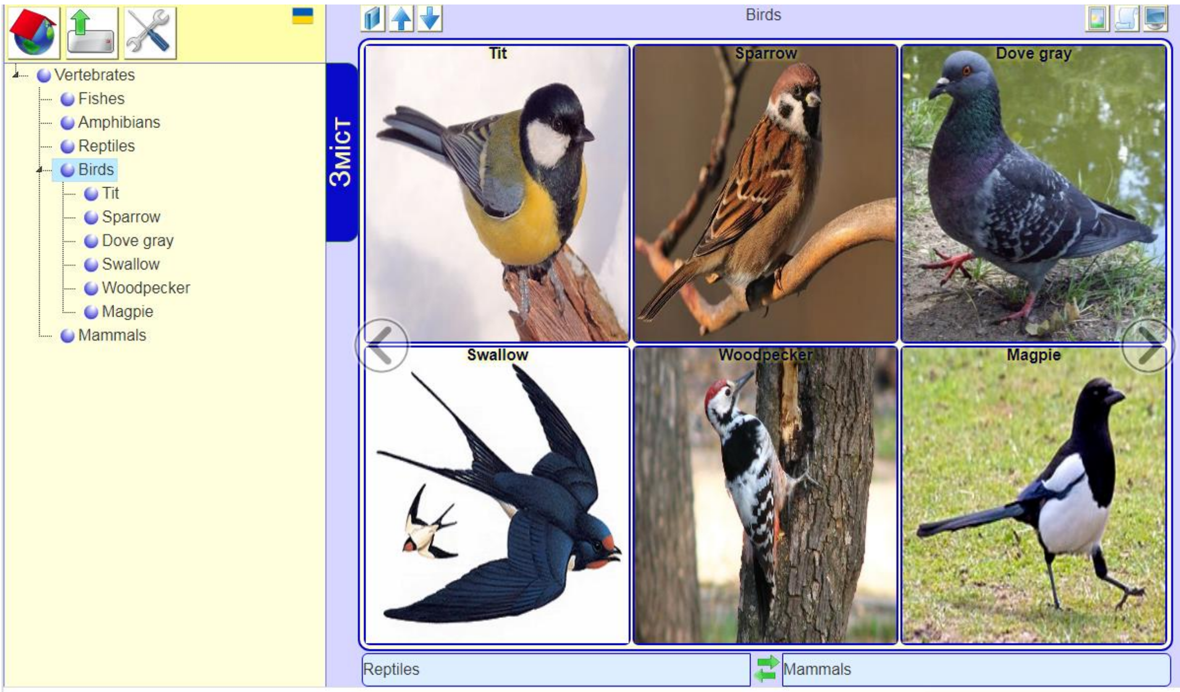

Fig. 3. Window ontograph nodes visualization

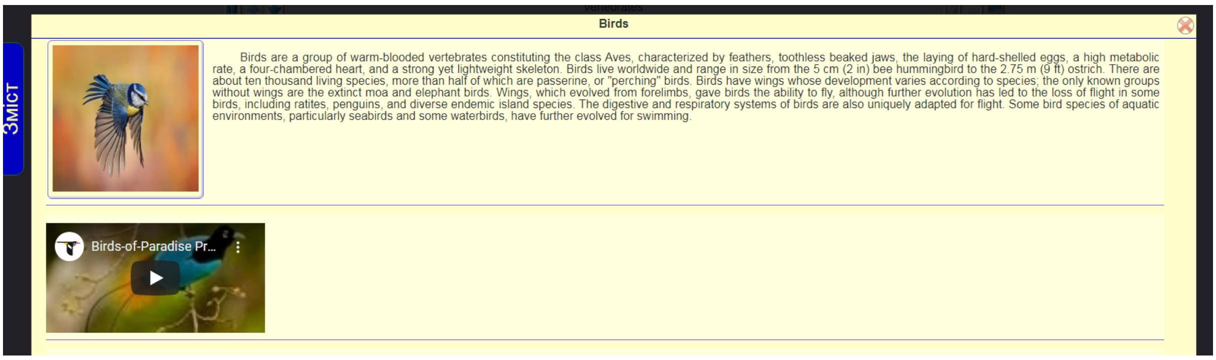

Fig. 4. Window visualization of the context of a single node

As page content, text, graphic images, sound and video files, documents in pdf, doc, docx, xls, presentations, web 2.0 applications and the like can be used.

5. Conclusions

The technology for visualizing an object ontology developed by the authors makes it possible to present the knowledge base in a form convenient for the end user. The software component responsible for viewing structured information supports two navigation systems: content with hyperlinks to transitions to individual nodes, which are displayed as a set of panels with contexts (as in electronic textbooks); navigation system with transitions to neighboring (single-level, parent and child) nodes, resembling a photo gallery slider navigator.

A key feature of the developed system is the simple implementation of contextual filling of the ontograph vertices. The data used can be represented by various types (plain text, image, sound, video, presentations, various document formats, including MS Office, etc.) allows us to provide detailed semantic annotation for collections of objects regardless of their nature.

We consider each of the ontological levels as a set of objects of the same type of visualization, united by spatial relationships. That allowed, applying the principle of “here and now”, significantly improve the usability of the user interface. In the pro-posed system, the user sees exactly the necessary interface elements and does not get distracted by those that are beyond the plane of the current task.

The ontological approach to servicing knowledge bases can be not only a means of organizing knowledge. Expanding the traditional functions of ontology software, you can create an information environment in which active work with knowledge is provided, and tasks related to education, the development of artificial intelligence, decision-making systems and many other areas where approximate sets can be used are solved in an original way.

References

[1] Voskresensky A. The concept of "Information": philosophical aspects (Analytical review)/ Anatoly Voskresensky // Theory and practice of social scientific information. 2013. No. 21. URL: https://cyberleninka.ru/article/n/ponyatie-informatsiya-filosofskie-aspekty-analiticheskiy-obzor

[2] Order of the Ministry of Education and Science, Youth and Sports of Ukraine 01.10.2012 № 1060 "On approval of the Regulations on electronic educational resources" // "Official Gazette of Ukraine", № 80 of October 29, 2012, p. 61–63

[3] Skaruk G. A. Electronic catalog as an object of quality assessment // Bibliosphere. 2018. № 2. P. 85–92. DOI: 10.20913/1815-3186-2018-2-85-92. URL: https://www.bibliosphere.ru/jour/article/view/242

[4] Guraliuk A. (2020). Library in the information society: an ontological approach / Andrii Guralyuk // Adaptive management: theory and practice. Series "Pedagogy". Vip. 8 (15). https://amtp.org.ua/index.php/journal/article/view/192/210

[5] Valiev R.M. Knowledge Bases in post-Industrial Society (2011). / Rustem Valiev //Space and time. 2011., no. 2 (4) URL: http://www.cawater-info.net/pdf/valiev.pdf

[6] Dorot V., Novikov F. Explanatory Dictionary of Modern Computer Vocabulary. St. Petersburg: BHV-Petersburg, 2004. 608 p. - ISBN 5-94157-491-6

[7] Ce Zhang, Christopher Ré, Cafarella Michael, Christopher De Sa, Alex Ratner, Jaeho Shin; Feiran Wang and Sen Wu (2017) *DeepDive: Declarative Knowledge Base Construction* Communications of the ACM. Vol. 60 Issue 5, p 93-102. 10 p. 7 Diagrams, 1 Graph. DOI: 10.1145/3060586 (EBSCO, Business Source Premier)

[8] Sathiya B, Geetha T V; Batyrshin Ildar Cross Valerie, Kreinovich Vladik and Rifqi Maria (2019) *A Review on Semantic Similarity Measures for Ontology.* Journal of Intelligent & Fuzzy Systems. Vol. 36 Issue 4, p 3045-3059. 15p. DOI: 10.3233/JIFS-18120 (EBSCO, Business Source Premier)

[9] Boskebeev K. Systematization of the Knowledge Base in Information Systems. / Kalychbek Boskebeev // Bulletin of TSUPBP. 2014. URL: http://www.cawater-info.net/pdf/boskebeev.pdf

[10] Karna, Nyoman & Supriana, Iping & Maulidevi, Nur. (2014). Intelligent Interface for Knowledge Based System. TELKOMNIKA. 12. 10.12928/telkomnika.v12i4.413.

[11] Nikonenko A.A. Overview of Ontological Knowledge Bases / A.A. Nikonenko // Artificial Intelligence. 2009. No. 4.S. 208-219.