Olga Yu. Gukosyants

Pyatigorsk State University,
9, pr. Kalinina, Pyatigorsk, 357532, Russia
gukosjants@mail.ru

Maxim N. Latu

Pyatigorsk State University,
9, pr. Kalinina, Pyatigorsk, 357532, Russia
laatuu@yandex.ru

Some features of query construction in an expert system according to the peculiarities of knowledge structuring (using the domain of nanotechnology as an example)

For citation: Gukosyants O. Yu., Latu M. N. Some features of query construction in an expert system according to the peculiarities of knowledge structuring (using the domain of nanotechnology as an example). *Vestnik of Saint Petersburg University. Language and Literature*. 2019, 16 (4): 638–652. https://doi.org/10.21638/spbu09.2019.406 (In Russian)

The article is devoted to the problem of information search intellectualization and knowledge extraction acceleration by means of specialized software — an expert system with scientific knowledge structured in the form of terminological networks. The objective of this research is to illustrate how to extract the necessary data from the expert system by stating certain requests and to find out what these requests are. The fragment of the terminological network of the terms of the category Instrument (artificial object) and adjacent terms reconstructed by the authors formed the basis of the article. The contexts with the technical terms of the considered category (*quantum dot, nanofibre, nanofilm, nanowire, nanoplate*) were the sources of the examples. Contextual analysis and the method of interpretation of the obtained data were used to work out semantic relations between the terms of the domain of nanotechnology. The allocated semantic relations made it possible to define the nature of the term system interaction. Due to the verification of the received semantic relations resulting in prototypical models for the technical terms mentioned above, the authors managed to reconstruct a fragment of the generalized terminological network. These findings served as the basis to identify some typical questions that should be used as search queries in the expert system. The article shows that the created search queries are common and enable users to obtain necessary information if the corresponding semantic relations in the terminological network of the required terms exist.

Keywords: technical term, expert system, systemic relation, prototypic semantic scheme, search query.

* The authors express their gratitude to the Ministry of science and higher education of the Russian Federation that financially supported the study: project No. 34.3234.2017/PCh "Development of linguistic principles of building an expert system for academic knowledge representation based on semantic terminological networks construction."

Introduction

The considerable volume of existing knowledge and its continuous expansion in different scientific areas have predetermined the importance of scientific knowledge ordering and systematization. One of the key purposes of this task is the construction of knowledge structures of certain domains not only for storage but also for fast data extraction from texts and information search intellectualization. Correspondingly, decision making based on an organized system of knowledge is one of the main tasks the current and developing knowledge-based systems are supposed to carry out.

The issue of knowledge systematization and modeling, a hierarchy of systemic relations between concepts, has a long history. The idea of the existence of a certain number of semantic groups to distinguish all lexical units was discussed even in the linguistic works of the 19th century (e. g.: [Pokrovskii 1895]). And the greatest resonance this theory received was with J. Trier's works [Trier 1931; 1934] and with the following generation of linguists [Shchur 1974; Apresian 1995; Karaulov 1972; Shaikevich 1963, etc.]. The term "matrix" in linguistic research has been used since 1935 (G. K. Zipf "The Psycho-Biology of Language: An Introduction to Dynamic Philology" [Zipf 1935]). The first semantic networks were applied by Silvio Ceccato [Ceccato 1961], Richard Richens [Richens 1956], Margaret Masterman [Masterman 1961] to define types of concept. And the first expert systems were DENDRAL (by Edward Feigenbaum in the 1960s [Feigenbaum, Buchanan 1978]) and MYCIN (by Edward Shortliffe in the 1970s [Shortliffe 1976]). According to Allen Newell, MYCIN was "a great-grandfather of expert systems" [Newell 1990].

Different ways and models of knowledge representation that exist these days are characterized by various principles of construction, their functional opportunities and solvable tasks. The analysis of some relevant methods of knowledge representation and the convenience of the implication of these methods in relation to the domain of nanotechnology have already been considered (in particular: [Gukosiants, Latu 2017; Latu 2018a]), and the advantages and disadvantages of some relevant methods of knowledge representation have been shown. Thus, the semantic (terminological) field organizes terminological units into groups according to certain signs or brightness of their manifestation, community of functions and combinations of lexical-syntactic signs, but does not reflect a term ratio with adjacent terms and links between them. Graphosemantic modeling reflects the analysis of the term's conceptual areas and represents a set of the term's conceptual fields and the links between them. Semantic graphs can also be complemented with components of the definitions inside the field and links between the components. Therefore, graphosemantic modeling does not demonstrate the position of the term in the ratio with adjacent terms, or their interaction, but gives an analysis of the conceptual areas of the term. Frame structures data of a certain subject domain, which has a ladder-type formation and assumes the existence of subsumption relations, reflecting the area of domain-specific knowledge. Frame characterizes a common cognitive context that distinguishes it from matrix, which is more sophisticated. Knowledge structuring in matrix is also aimed at the analysis of the concept structure, its signs and subject domains of its performance. Unlike frame, matrix does not characterize a common cognitive context but a system of contexts. Matrix can be based on a field principle as semantic graphs are. Like the semantic (terminological) field, matrix reflects signs of the analyzed objects. All these features make a matrix more suitable for the description of a concept, but do not make it possible to analyze terms. A terminological network represents a subtype of semantic networks that deals with academic knowledge expressed by technical terms only and its structure involves the semantic relations that may appear between academic concepts. Terminological network demonstrates the position of the term in a certain terminological system. Terminological network (as well as ontology) represents the relations between the system's components and is the most successful model of scientific knowledge structuring. Ontology is viewed as being more extensive (with non-terminological lexis) and context-dependent. Expert system is not just a method of knowledge representation but is a special computer program, which — by means of knowledge of a certain subject domain structured by domain experts — helps to find the solutions to a number of problems at the request of the user.

Along with various types of the knowledge-based systems an expert system is considered to be one of the most applicable for the scientific sphere. The main feature of the expert system is the knowledge base. The basic methods of knowledge presentation in expert system are "IF-THEN" rules. "Knowledge which consists only of logical relationships between facts, and which contains no element of doubt, is called categorical and can be expressed as 'IF-THEN' rules. By analogy with the term 'database', a collection of these rules is said to constitute a knowledge base" [Todd 1992: 25]. Thus, an expert system that uses a collection of facts, is based on extensive domain expert knowledge and is equipped with a certain number of sensible queries intends to serve as a consultant for decision making and makes knowledge available to a wider audience.

As R. S. Grabinger, B. W. Wilson and D. H. Jonassen (1990) state: "an expert system is a computer program that simulates the way human experts solve problems" [Grabinger et al. 1990: 2]. A similar definition of this concept can be found in Garcia et al. (2001): "a computer program that attempts to imitate the reasoning process and knowledge of experts in solving specific types of problems" [Garcia et al. 2001]. Therefore, an expert system is a computer program that can process large volumes of information and find answers to certain questions concerning a particular domain. What is important in this respect is how a user may interact with an expert system and retrieve the necessary data. Query-answer regime is one of the ways to do so [Alonso et al. 2012]. However the development of such tool may face a number of problems [Hu, Liu 2006] and new approaches to how this should be done and what principles should be taken into consideration are suggested and discussed [Sena, Furtado 1998].

The main objective of this research is to illustrate how knowledge is structured in an expert system and how to extract the necessary data by stating certain requests and what these requests are. The examples are drawn from the domain of nanotechnology and special attention is paid to the technical terms in the category Instrument, i. e the technical terms that determine all artificial objects and components of equipment/mechanism. The analysis of scientific papers was carried out and the fragments correlating with certain systemic relations between technical terms were given. The number of the studied text extracts is about 230 fragments.

Materials and methods

As I. M. Ahmed, A. M. Mahmoud, M. Aref and A.-B. M. Salem state, expert systems currently "solve part or whole of a significant problem" and "reduce the essential need for human experts" in the domains of "law, engineering, airspace, military, medicine, chemistry and banking" [Ahmed et al. 2013]. It is worth mentioning that one of the leading domains, where the expert systems are useful and helpful, is medicine; that is proven by a great number of papers that discuss their application (e. g.: [Ahmed et al. 2015; Singla et al. 2014; Zeki et al. 2012; Garcia et al. 2001; Kumar, Bhimrao 2012; Aliferis, Miller 1995; Arsene et al. 2015, Shabut et al. 2018]). Moreover, even educational sphere and quality of life evaluations can be regulated by expert systems [Lee et al. 2018; Herrero-Jiménez 2012; Atanasova, Krupka 2013].

Talking about the architecture of an expert system, three main software components, which I. M. Ahmed et al. describe, should not go unmentioned: "(1) knowledge base (2) inference engine and (3) user interface" [Ahmed et al. 2015]. In this paper we shall focus on the system of concepts and systemic relations between them that appear in the knowledge database, the requests a user can state in the user interface, and how one can use the correlation between certain requests and systemic relations to extract the necessary data. An expert system, as a powerful tool to provide information and solve a complex problem of the user, should be fitted with a broad database. The knowledge in the database of an expert system can be accumulated from experts and professionals themselves, as well as various academic texts — articles, books, dictionaries and other related sources.

As previous research has demonstrated [Fidel, Efthimiadis 1995; Johannsen, Alty 1991; Potapova, Shirokov 2013; Gukosiants, Latu 2017; Latu 2018a], a terminological semantic network represents a model of knowledge organization on which an expert system can be based. The vertices of the network are the domain concepts and the technical terms that express them; the links between the vertices demonstrate semantic relations between them [Malkovskii, Solov'ev 2012]. The arrangement of concepts by means of the terminological network allows the position of the concept in the system of knowledge to be defined, as well as the adjacent terms and the different characters of their relationships to be identified. As we discuss further on, the latter is crucial for query construction in an expert system.

The technical terms *quantum dot, nanofibre, nanofilm, nanowire* and *nanoplate* and their relations with the adjacent technical terms are considered in the article; their charts are taken as bases.

The principles of terminological network construction, the specificity of adjacent and non-adjacent terms and the character of their associativity is brought up by M. N. Latu and A. A. Levit [Latu, Levit 2016]. The construction of terminological knowledge bases is described by P. Faber et al. who says that the bases are "based on an underlying network of semantic relations, and make an effort to encode information in the form of knowledge" and "semantic relations largely depend on the type of entity being described, its nature, and relational power" [Faber et al. 2009].

During the terminological network building we recognize that it is possible to differentiate the terms of the nanotechnological sphere according to the designated concept categories, as academic concepts may represent material or non-material entities, objects of natural or artificial nature. The categories with a material referent include: Natural object/phenomenon, Substance, Locus, Instrument/product, Mechanism, Material and Construction. The categories with a non-material referent are: Agent, Characteristic, Ideal phenomenon, Situation and Process [Latu 2015].

The system of terms involves a complex system of the semantic relations between them. E. Wüster and D. S. Lotte are considered the founders of the study of terms; their first works appeared in the early 1930s [Wüster 1931; Lotte 1931]. E. Wüster created the "traditional" classification of systemic relations. He divided the relations between the terms into ontological (the contiguity relation in space and time) and logical (the similarity relation).

One of the basic hierarchies of the relations between terms of qualitative and quantitative relation categories was offered by I. Dahlberg [Dahlberg 1974; 2014]. Categorization of professional terms (the terms of processes, objects, properties, sciences, professions, etc.) and the identification of the relations between these terms can be found in T. L. Kandelaki's works [Kandelaki 1977].

Despite the lasting history of the systemic relations development, the information search organization and the search queries analysis (the main functions of the developing expert system) demand reconsideration of the existing types of systemic relations, their supplement and specification.

As for the systemic relations, M. G. Malkovsky and S. Y. Soloviev talk about the existence of two types of binary semantic relation between the terms in terminological networks — "is a" and "related to" relations [Malkovskii, Solov'ev 2012]. In our research we rely on the following types of semantic relation that link the technical terms in a terminological network:

- AKO — "a kind of," a link between a generic term and its direct hyponyms;
- ISA — "is a," a relation of coincidence or inclusion in a set;
- PO — "part of," the relation between a part and the whole;
- At — "attribute," the relation between a property and the referent it characterizes;
- S — "subject," the link between the initiator of the process and the process itself;
- Obj — "object," the link between the object under influence and the process;
- Loc — "location," the link between a referent and its location;
- Sr — "source," the link between the initial position of a referent and the process in which it is involved;
- Rec — "recipient," the link between the final position of a referent and the process in which it is involved;
- Inst — "instrument," the link between the instrument used in the process and the process itself;
- R — "result," the relation between a process and its result, etc. [Latu, Levit 2016].

Our choice of this particular body of semantic relations for the study and the construction of expert systems is predetermined by the fact that our earlier research has proved them to be the most frequent and productive systemic links that exist between the discussed categories of academic concepts [Latu, Levit 2017]. The comprehensive body of these highly productive types of systemic relations was listed and their peculiarities were discussed [Latu 2018b]. It is worth mentioning in this respect that these defined systemic relations appear to be universal (since the research encompassed various fields of academic knowledge) and not specific to particular domains only. The latter represents one of the common problems since very often different scholars define and focus on the relations specific to a particular field of knowledge that are not applicable to the others and are not productive in general. Another common difficulty is that one and the same type of systemic relations may be expressed by different terms in various works. However one should also bear in mind that their understanding and descriptions may also slightly differ. Thus, for example, Ph. Mayr and V. Petras and B. Lauser et al. mention

"Equivalence (=) means identity, synonym, quasi-synonym; Hierarchy (Broader terms <; narrower terms >); Association (^) for related terms; An exception is the Null (0) relation, which means that a term can't be mapped to another term" [Mayr, Petras 2008; Lauser et al. 2008]. S. Rirdance and A. Vasiļjevs describe the concept of relations and divide them into generic, partitive and associative ones [Rirdance, Vasiļjevs 2006]. P. Faber et al., who describe the terminological knowledge base in environmental engineering, differentiate the following conceptual relations:

> "ISA — this generic-specific relation reflects hierarchical inheritance in the conceptual network of the domain; PART-OF — this relation also reflects the hierarchical structure of the domain; this relation directly refers to the parts of each concept; MADE-OF — this relation links both artificial and natural objects to the material they are made of, and thus bears a certain resemblance to the PART-OF relation; DELIMITED-BY — this relation is used for physical objects, and marks the boundaries, dividing one object from another; LOCATED-AT — this relation is relevant when the location of a physical object is an essential characteristic for its description; TAKES-PLACE-IN — this relation describes the context of processes which have spatial and temporal dimensions; ATTRIBUTE-OF — this relation is only useful for concepts designated by specialized adjectives, such as isotropic, alluvial, abyssal, etc., or nouns that designate the properties of other concepts, such as altitude, capacity, coefficient, etc.; RESULT-OF — this relation is relevant to either processes or entities that are derived from other processes; AFFECTS — this relation links processes or objects that cause a change in any other object or process without producing a final result; HAS-FUNCTION — this relation not only links objects or processes that are artificially created or carried out with a specific function, but also natural entities which, despite not being goal-directed, can be used for human profit; EFFECTED-BY — this relation is only used for instruments that carry out some process or create an entity" [Faber et al. 2009; Garcia, Faber 2017].

As one can easily see, the number and specificity of the systemic relations within different approaches to terminological network construction have both similarities and distinctions.

Results

This work deals with the technical terms of the category Instrument and their systemic relations with the adjacent terms of various other categories. As noted above, a terminological semantic network serves as the foundation for an expert system. The significant elements of the network are its vertices (the domain concepts and the technical terms) and the links between the vertices (the semantic relations between the terms). Therefore, it is possible to allocate more complex patterns in the network — the schemes of two adjacent terms correlation. The structure of such prototypic schemes includes four basic elements: Category 1 (the first term), Category 2 (the second term), the system relation between the terms and the direction of the vector of system relation [Latu 2015; 2018b].

To identify the prototypic semantic schemes of the category Instrument, the terms of this category were extracted from special dictionaries, official standards and an encyclopedia of nanotechnology. After that, scientific and technical texts were analyzed and the adjacent terms that are linked by different systemic relations to the considered technological terms of the category of Instrument were collected. The extracted data was represented in the charts with the considered technological terms of the category of Instrument in the centre and the adjacent terms around it connected to each other by targeting vectors demonstrating the semantic relations.

After the analysis of scientific and technical texts to identify the semantic schemes of the technical terms *quantum dot, nanofibre, nanofilm, nanowire* and *nanoplate*, some prototypic schemes that are equal for all these terms were distinguished. These prototypic schemes are presented in a fragment of the terminological network that reflects the systemic relations of the nanotechnological terms in the category Instrument (Fig.).

An expert system based on such a diagram, with its vertices and links of different types between them, enables users to reach necessary information just by typing in a search query. To do so, the graphical user interface of the expert system has to be equipped with a special dialog box for users to type in a query construction dialog or to select the query from the given variants. And to get them and to make the queries sensible for the expert system, its engineers have to create a base of common questions according to the revealed prototypic schemes in the expert system.

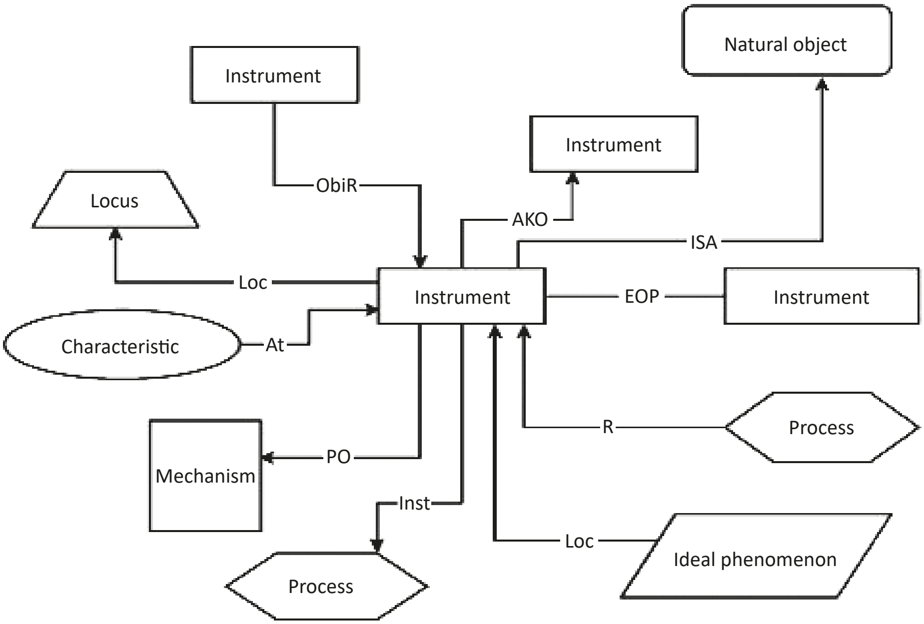

Fig. A fragment of the terminological network that reflects the systemic relations of the nanotechnological terms of the category Instrument

Taking into account the systemic relations in the fragment of the terminological network reflecting the systemic relations of the nanotechnological terms of the category Instrument (Fig.), we now introduce some typical questions that should be used as search queries in the expert system. Presuming that each of the systemic relations mentioned below can be expressed by means of a certain number of typical questions, we are going to define the most frequent search queries with the help of which the user can exclude the data from the expert system.

1. One of the basic relations that all Instruments and Objects have is AKO ("a kind of"). Thus, there exists a prototypic pattern: **Instrument — AKO → Instrument**. The questions that correlate with this systemic relation are: *What is _____? What is the generic term of _____? What kinds does _____ have? What does _____ fall into?*

The following examples of the AKO relation were discovered in the articles devoted to nanofibers.

(1) Small **quantum dots**, *such as* **colloidal semiconductor nanocrystals**, can be as small as 2 to 10 nanometers, corresponding to 10 to 50 atoms in diameter and a total of 100 to 100,000 atoms within the quantum dot volume [Kumar 2015].

The relation between the hyponym and its generic term is visible in the following statement where "such as" serves as a linguistic marker for this systematic relation. This lexical means points to the existence of some types of quantum dot, one of which is a colloidal semiconductor nanocrystal.

2. The prototypic scheme **Instrument — ISA → Natural object** correlates with the queries: *How _____ is treated*? *How _____ is regarded*?

(2) ***Nanofiber* — *nano-object*** *(flexible or rigid) with two external dimensions in the nanoscale and the third dimension significantly larger* [Trotta, Mele 2019].

The given context reflects the relation between the terms "nano-object" and "nanofiber." The fact that a "nanofiber" is defined as a "nano-object" proves the inclusion of nanofibers in a set of nano-objects without specifying the definite group of nano-objects. Such a relation can undoubtedly be converted into one of the basic semantic relations between the terms — a relation of coincidence or inclusion in a group (ISA).

This fragment correlates with one of the questions we determined above as a typical one for the prototypic scheme **Instrument — ISA → Natural object.**

3. The search query for the relation **EOp** (the relation that exists between the hyponyms of the same generic term) can be expressed with the questions: *What concept is adjacent to _____? What is the adjacent term to _____? What terms are adjacent to ____?*

(3) **Nanowires, nanobelts, nanoribbons, nanorods** are a new class of **quasi-one-dimensional materials** that have been attracting a great research interest in the last few years [Wang 2003]. The phrase "are a new class of" serves as linguistic marker to unite the terms "nanowires," "nanobelts," "nanoribbons" and "nanorods" as "quasi-one-dimensional materials" and identifies all of these terms as the hyponyms of one generic term "quasi-one-dimensional materials."

(4) ***Nano-objects*** *(*for example*, **nanoparticles, nanofibres,** and **nanoplates**), often occur in (large) groups, rather than isolated*[1]. In the given example the generic term "nano-object" correlates with its hyponyms, located behind the main term in brackets after the phrase "for example." This phrase demonstrates the involvement of all terms in brackets into one group and their equal position inside the group that makes them adjacent concepts to each other.

[1] State Standard ISO/TS 27687:2008(en): Nanotechnologies — Terminology and definitions for nano-objects — Nanoparticle, nanofibre and nanoplate. https://www.iso.org/obp/ui/ iso:std:iso:ts:27687:ed-1:v2:en:sec:4 (accessed date: 10.03.2019).

4. To identify the basic characteristics of the analyzed referents of the category Instrument the following questions can be typed in as requests: *What characteristics does _____ have? What properties does _____ possess? What properties does _____ exhibit?*

The confirmation of the validity of these questions could be analyzed with a focus on how terms are used in texts. So, the semantic relations between the terms are extracted from the knowledge-rich contexts.

(5) **Electrical properties** *of* individual **carbon nanofibers** are statistically determined by a current sensing atomic force microscopy mode [Fourdrinier et al. 2008].

(6) Some **carbon nanofibers** *exhibit* very low **resistances** (few kilohms), implying that good contact is obtained between the nanofiber and the substrate, while others carbon nanofibers exhibit high resistance, attributed to local poor electrical contacts between carbon nanofibers and TiN layer [Ibid.].

(7) The **nanofilm** *is assumed to be* **elastic** with Young's modulus E=2GPa [Peng 2015: 29].

The underlined phrases serve as markers to identify the features and properties of the technical terms.

5. To clarify the processes that lead to Instrument creation **(Instrument ← R — Process**) the following questions can be typed in as requests: *How is _____ created? What process resulted in _____? What method is used to make/prepare/synthesize _____? What are the synthesis techniques for _____? What kind of techniques are available for the synthesis of _____?*

The prototypic scheme between the categories Instrument and Process is represented in the following fragments.

(8) **Electrodeposition** *can be used to prepare* metal chalcogenide **nanostructured film** [Sharma 2015].

(9) **Quantum dots** have attracted most interest because of *their* interesting **optical properties** [Woodford 2018].

(10) Currently, there are three *techniques* available *for the synthesis of* **nanofibres: electrospinning, self-assembly, and phase separation** [Vasita, Katti 2006].

6. A fragment of the terminological network that reflects systemic relations of the nanotechnological terms of the category Instrument (Fig.) demonstrates one more prototypic scheme — the semantic relation of the instruments that are used to carry out some processes (**Instrument — Inst → Process**).

The questions for this scheme are: *What _____ is used in the process _____? What _____ is used for _____?* (11) *Composite structures are often the way to support **nanofiber*** for *high efficiency* ***filtration*** [Hutten 2007]. The given example identifies the application of nanofiber in the process of filtration.

7. Another very productive scheme is: ***Instrument ← Loc — Ideal phenomenon***, which demonstrates an abstract phenomenon that takes place in or on the referent of the category Instrument. The questions that correlate with the Loc systemic relation are: *What phenomenon occurs in _____? What phenomenon is related to _____?*

(11) The first energy system assumes that the **nanowire** heterostructure is coherently strained or *contains* one misfit **dislocation** [Kuech 2014: 424].

(12) Using particle in a box model, we have studied the **quantum confinement effects** *on* **quantum dots** (QDs) [Woodford 2018].

Both "dislocation" and "quantum confinement effects" are not material objects, but abstract mental images that are the results of deduction; they refer to the category Ideal phenomenon. Thus, the verb "contains" characterizes the presence of misfit dislocation in the nanowire, and the preposition "on" characterizes the location of quantum confinement effects on quantum dots.

8. One more prototypic scheme demonstrated in Fig. is: **Instrument — PO → Mechanism**, which means that artificial tools or instruments could be components of sophisticated mechanisms. The query for this systemic relation is: *What are the parts of _____? What does _____ consist of?*

Artificial tools as integral parts of mechanisms appear in the following statements.

(13) In one simple and relatively trivial application, a thin **filter** *made of* **quantum dots** has been developed so *it can be fitted on* top of a **fluorescent or LED lamp** and convert its light from a blueish color to a warmer, redder, more attractive shade similar to the light produced by old-fashioned incandescent lamps [Ibid.].

The phrase "made of" reflects the partitive relation between quantum dots and thin filters, and fluorescent or LED lamps. As the constituent parts of thin filters, quantum dots can be identified as the components of fluorescent or LED lamps as well. Both thin filters and fluorescent or LED lamps are complicated mechanisms with compound internal designs.

The same relation is traced between "quantum dots" and "computer screens" and "displays" in the following fragment. (15) ***Quantum dots*** *are also finding their way* into ***computer screens*** *and* ***displays****, where they offer three important advantages* [Ibid.]. The list of prototypic schemes of the terms of the category Instrument is not limited to the items we have discussed here. We cannot help mentioning the following less common prototypic schemes, which are not demonstrated in Fig.:

- ***Instrument — Loc → Locus:*** *Where could _____ be found?* (when the user is interested in the place in which the Instrument is situated);
- ***Instrument ← ObjR — Instrument:*** *What is _____ produced by?* (when the user is interested in another Instrument that was used to create the Instrument).

Filling the gaps in the questions with the terms of the category Instrument the user can exclude all the data that the system contains.

Conclusions

We have presented the first systematic study of the expert system of the domain of nanotechnology as an important tool of knowledge representation and visualization, achieving accuracy on large data sets acquired from texts, standards, encyclopedias and other related sources.

We applied some methods from discourse analysis and contextual methods to transformational analysis and statistical methods. The prototypic schemes we have talked about are similar across different data sets with the same fragments and links.

Undoubtedly, the study has not been finished yet and is aimed at the extraction of all possible prototypic schemes of all the mentioned categories of terms existing in the domain of nanotechnology.

Sources

Fourdrinier et al. 2008 — Fourdrinier L., Le Poche H., Chevalier N., Mariolle D., Rouviere E. Electrical properties measurements on individual carbon nanofibers by scanning spreading resistance microscopy. *Journal of Applied Physics*. 2008, 104 (11). DOI: https://doi.org/10.1063/1.3033491.

Hutten 2007 — Hutten I. M. *Handbook of Nonwoven Filter Media*. Oxford; Burlington (MA): Butterworth-Heinemann, 2007. 496 p.

Kuech 2014 — *Handbook of Crystal Growth*: in 3 vols. 2nd ed. Vol. 3: Thin Films and Epitaxy. Kuech T. (ed.). [S. l.]: Elsevier, 2014. 1382 p.

Kumar 2015 — Kumar A. *Introduction to solid state physics*. 2nd ed. Delhi: PHI Learning Private Limited, 2015. 528 p. https://books.google.ru/books?id=b_-UCgAAQBAJ&pg (access date: 25.02.2019).

Peng 2015 — Peng Z. *Bio-inspired studies on adhesion of a thin film on a rigid substrate*. Heidelberg: Springer, 2015. 97 p.

Sharma 2015 — Sharma K. R. Process considerations for nanostructured coatings. In: *Anti-abrasive nanocoatings: current and future applications*. Aliofkhazraei M. (ed.). Amsterdam: Elsevier: Woodhead Publishing, 2015. P. 137–153. DOI: https://doi.org/10.1016/B978-0-85709-211-3.00006-6.

Trotta, Mele 2019 — Trotta F., Mele A. *Nanosponges: synthesis and applications*. [S. l.]: John Wiley & Sons, 2019. 350 p. https://books.google.ru/books?id=DteFDwAAQBAJ&pg (access date: 10.03.2019).

Vasita, Katti 2006 — Vasita R., Katti D. S. Nanofibers and their applications in tissue engineering. *International Journal of Nanomedicine*. 2006, l (1): 15–30. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2426767/ (access date: 01.03.2019).

Wang 2003 — Wang Z. L. *Nanowires and nanobelts: materials, properties and devices*. Vol. 1: Metal and aemiconductor nanowires. New York; London: Springer, 2003. 475 p. DOI: https://doi.org/10.1007/978-0-387-28745-4.

Woodford 2018 — Woodford C. What are quantum dots? *Explain that stuff. The FREE online science and technology book*. 2012, updated: July 28, 2018. https://www.explainthatstuff.com/quantum-dots.html (access date: 20.03.2019).

References

Ahmed et al. 2013 — Ahmed I. M., Alfonse M., Aref M., Salem A.-B. M. Reasoning techniques for diabetics expert systems. *Procedia Computer Science*. 2015, 65: 813–820. DOI: https://doi.org/10.1016/j.procs.2015.09.030.

Ahmed et al. 2015 — Ahmed I. M., Mahmoud A. M., Aref M., Salem A.-B. M. A study on expert systems for diabetic diagnosis and treatment. In: *Recent advances in information science.* Proceedings of the 7th European computing conference (ECC ’13), Dubrovnik, Croatia, June 25–27, 2013. Boras D., Damij N. (eds.). Greece: WSEAS Press, 2013. P. 363–367. https://pdfs.semanticscholar.org/f090/0fd503b058219cec612cbe770cb03c243d0a.pdf (access date: 20.01.2019).

Aliferis, Miller 1995 — Aliferis C. F., Miller R. A. On the heuristic nature of medical decision-support systems. *Methods of Information in Medicine.* 1995, 34: 5–14. DOI: https://doi.org/10.1055/s-0038-1634584.

Alonso et al. 2012 — Alonso O., De La Pena P., Moreno A. Data mining and query answer techniques applied to a bio-nutritional trials focused expert system. *Artificial Intelligence Research and Development*. 2012, 248: 9–18.

Apresian 1995 — Apresian Iu. D. *Selected compositions:* in 2 vols. Vol. 1: Lexical semantics. Moscow: Iazyki russkoi kul’tury Publ., 1995. 472 p. (In Russian)

Arsene et al. 2015 — Arsene O., Dumitrache I., Mihu I. Expert system for medicine diagnosis using software agents. *Expert Systems with Applications*. 2015, 42 (4): 1825–1834. DOI: https://doi.org/10.1016/j.eswa.2014.10.026.

Atanasova, Krupka 2013 — Atanasova I., Krupka J. Architecture and design of expert system for quality of life evaluation. *Informatica Economica*. 2013, 17 (3): 28–35. DOI: https://doi.org/10.12948/issn14531305/17.3.2013.03.

Ceccato 1961 — Ceccato S. *Linguistic analysis and programming for mechanical translation*. Milano: Feltrinelli Editore, 1961. 242 p.

Dahlberg 1974 — Dahlberg I. *Grundlagen universaler Wissensordnung: Probleme und Möglichkeiten eines universalen Klassifikationssystems des Wissens*. München: Verlag Dokumentation, 1974. 366 S.

Dahlberg 2014 — Dahlberg I. *Wissensorganisation. Entwicklung. Aufgabe. Anwendung. Zukunft*. Würzburg: Ergjn Verlag, 2014. 174 S.

Faber et al. 2009 — Faber P., Leon P., Prieto J. Semantic relations, dynamicity, and terminological knowledge bases. *Current Issues in Language Studies*. 2009, 1: 1–23. https://www.researchgate.net/publication/235743149_Semantic_relations_dynamicity_and_terminological_knowledge_bases (access date: 16.12.2018).

Feigenbaum, Buchanan 1978 — Feigenbaum E., Buchanan B. Dendral and meta-dendral. *Artificial Intelligence*. 1978, 11: 5–24.

Fidel, Efthimiadis 1995 — Fidel R, Efthimiadis E. N. Terminological knowledge structure for intermediary expert Systems. *Information Processing & Management*. 1995, 31 (1): 15–27. DOI: https://doi.org/10.1016/0306-4573(94)E0004-L.

Garcia et al. 2001 — Garcia M. A., Gandhi A. J., Singh T., Duarte L., Shen R., Dantu M., Ponder S., Ramirez H. Esdiabetes (an expert system in diabetes). *Journal of Computing Sciences in Colleges*. 2001, 16 (3): 166–175. https://dl.acm.org/citation.cfm?id=374762 (access date: 28.01.2019).

Garcia, Faber 2017 — Garcia M. C., Faber P. A semantic approach to the inclusion of complex nominals in english terminographic resources. In: *Computational and corpus-based phraseology: Second International Conference, Europhras 2017, London, UK, November 13–14, 2017, proceedings*. Mitkov R. (ed.). Cham (Switzerland): Springer, 2017. P. 145–159. DOI: https://doi.org/10.1007/978-3-319-69805-2_11.

Grabinger et al. 1990 — Grabinger R. S., Wilson B. W., Jonassen D. H. *Building expert systems in training and education*. New York; London: Praeger, 1990. 173 p.

Gukosiants, Latu 2017 — Gukosiants O. Iu., Latu M. N. Some ways of studying and modeling of scientific knowledge representation. *Philology: scientific research*. 2017, 4: 32–49. DOI: https://doi.org/10.7256/2454-0749.2017.4.24141. (In Russian)

Herrero-Jiménez 2012 — Herrero-Jiménez C. M. An expert system for the identification of environmental impact based on a geographic information system. *Expert Systems with Applications*. 2012, 39 (8): 6672–6682. DOI: https://doi.org/10.1016/j.eswa.2011.10.019.

Hu, Liu 2006 — Hu J., Liu Y. A rough set based approach for solving the unmatching problem between user query and rules inexpert system. *The 5th Wuhan International Conference on E-Business, China University of Geo-Sciences, Wuhan, Peoples Republic of China, 27–28 May 2006. Proceedings*. Duserick F. G. (ed.). Wuhan, 2006. Vol. 1. P. 238–243.

Johannsen, Alty 1991 — Johannsen G., Alty J. Knowledge engineering for industrial experts systems. *Automatica*. 1991, 27 (1): 97–114. DOI: https://doi.org/10.1016/0005-1098(91)90009-Q.

Kandelaki 1977 — Kandelaki T. L. *Semantics and motivation of terms*. Moscow: Nauka Publ., 1977. 167 p. (In Russian)

Karaulov 1972 — Karaulov Iu. N. The structure of lexical-semantic dield. *Philological sciences*. 1972, 1: 57–68. (In Russian)

Kumar, Bhimrao 2012 — Kumar S., Bhimrao B. Development of knowledge base expert system for natural treatment of diabetes disease. *International Journal of Advanced Computer Science and Applications*. 2012, 3 (3): 44–47. DOI: https://doi.org/10.14569/IJACSA.2012.030308.

Latu 2015 — Latu M. N. Peculiarities of logical categories representation in the underlying forms of technical terms. *Issues of Cognitive Linguistics*. 2015, 2 (43): 100–106. https://elibrary.ru/item.asp?id=23286741 (access date: 06.02.2019). (In Russian)

Latu 2018a — Latu M. N. Prototypical models of academic concepts correlation in a terminological network. *Cognitive Studies of Language*. 2018, 34: 710–713. https://elibrary.ru/item.asp?id=36423843 (access date: 08.02.2019). (In Russian)

Latu 2018b — Latu M. N. Types of systemic relations between technical terms in the academic knowledge network models. *Issues of Cognitive Linguistics*. 2018, 4 (57): 134–142URL: https://elibrary.ru/item.asp?id=35670033 (access date: 06.02.2019). (In Russian)

Latu, Levit 2016 — Latu M. N., Levit A. A. The study of systemacy of technical terms by means of semantic network. *Bulletin of the Kalmyk Institute for Humanities of the Russian Academy of Sciences*. 2016, 2 (24): 116–124. https://elibrary.ru/item.asp?id=27289972 (access date: 06.02.2019). (In Russian)

Latu, Levit 2017 — Latu M., Levit A. Technical terms of the category Locus and their semantic relations within a terminology. *The Rupkatha Journal on Interdisciplinary Studies in Humanities*. 2017, 9 (1): 21–35.

Lauser et al. 2008 — Lauser B., Johannsen G., Caracciolo C., van Hage W. R., Keizer J., Mayr P., Greenberg J. Comparing Human and Automatic Thesaurus Mapping Approaches in the Agricultural Domain. In: *Metadata for semantic and social applications: proceedings of the International Conference on Dublin Core and Metadata Applications. Berlin, 22–26 September 2008*. Greenberg J., Klas W. (eds.). Singapore: Dublin Core Metadata Initiative; Göttingen: Universitätsverlag Göttingen, 2008. P. 43–53. DOI: http://dx.doi.org/10.18452/1251. http://edoc.hu-berlin.de/18452/1903 (access date: 14.01.2019).

Lee et al. 2018 — Lee W.-K., Leong C.-F., Lai W.-K., Leow L. K., Yap T.-H. ArchCam: Real time expert system for suspicious behavior detection in ATM site. *Expert Systems with Applications*. 2018, 109 (1): 12–24. DOI: https://doi.org/10.1016/j.eswa.2018.05.014.

Lotte 1931 — Lotte D. S. Immediate Tasksoftechnical Terminology. *Izvestia of the USSR Academy of Sciences. Department of Social Sciences*. Series 7. 1931, 4: 883–891. (In Russian)

Malkovskii, Solov'ev 2012 — Mal'kovskii M. G., Solov'ev S. Iu. Terminological networks. *Open Semantic Technologies of Intellectual Systems Projection*. 2012, 2: 77–82. https://elibrary.ru/item.asp?id=32704351. (In Russian)

Masterman 1961 — Masterman M. Semantic message detection for machine translation, using an interlingua. In: *International Conference on Machine Translation of Languages and Applied Language Analysis. National Physical Laboratory, Teddington, UK, 5–8 September 1961: proceedings*. London: H. M. S. O., 1962. P. 438–475

Mayr, Petras 2008 — Mayr P., Petras V. Building a terminology network for search: the KoMoHe project. In: *Metadata for semantic and social applications: proceedings of the International Conference on Dublin Core and Metadata Applications. Berlin, 22–26 September 2008*. Greenberg J., Klas W. (eds.). Singapore: Dublin Core Metadata Initiative; Göttingen: Universitätsverlag Göttingen, 2008. P. 177–182. DOI: http://dx.doi.org/10.18452/1264. http://edoc.hu-berlin.de/18452/1916 (access date: 05.12.2018).

Newell 1990 — Newell A. *Unified theories of cognition*. Cambridge (MA): Harvard University Press, 1990. xvii, 549 p.

Pokrovskii 1895 — Pokrovskii M. M. Semasiological research in the field of ancient language. Moscow: University Press, 1895. 131 p. (In Russian)

Potapova, Shirokov 2013 — Potapova E. V., Shirokov V. A. Multilingual ontology modeling as lexicographic system. *Artificial intelligence and decision making*. 2013, 2: 23–31. https://elibrary.ru/item.asp?id=19112892 (access date: 05.12.2018). (In Russian)

Richens 1956 — Richens R. H. Preprogramming for mechanical translation. *Mechanical Translation*. 1956, 3 (1): 20–25.

Rirdance, Vasiļjevs 2006 — Rirdance S., Vasiļjevs A. *Towards consolidation of european terminology resources*. Riga: Tilde, 2006. 123 p.

Sena, Furtado 1998 — Sena G. F. de, Furtado A. L. An extended query-the-user tool for expert systems. In: *Proceedings of the 4th World Congress of Expert Systems: applications of advanced information technologies; 1998; Mexico City.* Vol. 1. New York: Cognizant Communication, 1998. P. 355–362.

Shabut et al. 2018 — Shabut A. M., Tania M. H., Lwin K. T., Evans B. A., Yusof N. A., Abu-Hassan K. J., Hossain M. A. An intelligent mobile-enabled expert system for tuberculosis disease diagnosis in real time. *Expert Systems with Applications*. 2018, 114: 65–77. DOI: https://doi.org/10.1016/j.eswa.2018.07.014.

Shaikevich 1963 — Shaikevich A. Ia. Distribution of words in the text and selection of semantic fields. *Foreign Language at School.* 1963, 2: 14–26. (In Russian)

Shchur 1974 — Shchur G. S. *The theory of a field in linguistics*. Moscow: Nauka Publ., 1974. 256 p. (In Russian)
Shortliffe 1976 — Shortliffe E. *Computer based medical consultations: MYCIN*. New York; Oxford: Elsevier, 1976. 286 p.
Singla et al. 2014 — Singla J., Grover D., Bhandari A. Medical expert systems for diagnosis of various diseases. *International Journal of Computer Applications*. 2014, 93 (7): 36–43. DOI: https://doi.org/10.5120/16230-5717.
Todd 1992 — Todd B. S. *An introduction to expert systems*. Oxford: University of Oxford, Computer Laboratory Programming Research Group, 1992. 95 p.
Trier 1931 — Trier J. *Der deutsche Wortschatz im Sinnbezirk des Verstandes*. Heidelberg: Winter, 1931. 347 S.
Trier 1934 — Trier J. Das sprachliche Feld. Eine Auseinandersetzung. *Neue Jahrbücher für Wissenschaft und Jugendbildung*. 1934, 10: 428–449.
Wüster 1931 — Wüster E. *Internationale Sprachnormung in der Technik. Besonders in der Elektrotechnik (Die nationale Sprachnormung und ihre Verallgemeinerung)*. Berlin: VDI-Verlag, 1931. 191 S.
Zeki et al. 2012 — Zeki T. S., Malakooti M. V., Ataeipoor Y., Tabibi S. T. An expert system for diabetes diagnosis. *American Academic & Scholarly Research Journal*. 2012, 4 (5): 1–13. https://www.questia.com/library/journal/1P3-2995404381/an-expert-system-for-diabetes-diagnosis (access date: 16.12.2018).
Zipf 1935 — Zipf G. K. *The psycho-biology of language: An introduction to dynamic philology*. Boston: The Riverside Press Cambridge, 1935. 336 p.

Received: May 30, 2019
Accepted: September 23, 2019

Гукосьянц Ольга Юрьевна

Пятигорский государственный университет,
Россия, 357532, Пятигорск, пр. Калинина, 9
gukosjants@mail.ru

Лату Максим Николаевич

Пятигорский государственный университет,
Россия, 357532, Пятигорск, пр. Калинина, 9
laatuu@yandex.ru

Принципы построения вопросов к экспертной системе в соответствии с особенностями организации знания (на примере сферы нанотехнологий)

Для цитирования: Гукосьянц О. Ю., Лату М. Н. Принципы построения вопросов к экспертной системе в соответствии с особенностями организации знания (на примере сферы нанотехнологий). *Вестник Санкт-Петербургского университета. Язык и литература*. 2019, 16 (4): 638–652. https://doi.org/10.21638/spbu09.2019.406

Статья посвящена вопросу интеллектуализации информационного поиска и ускорения процесса извлечения научных знаний из специализированного программного обеспечения — экспертной системы, научные знания в которой структурированы в виде терминологических сетей. Задачей исследования явилось представление способа извлечения необходимой информации из экспертной системы путем ввода конкретного запроса и перечисление возможных запросов для системы. Базовой терминологической

* Публикация выполнена в рамках реализации проекта «Разработка лингвистических принципов проектирования и создание экспертной системы представления элементов научного знания на основе конструирования терминологических семантических сетей» по проектной части государственного задания Министерства образования и науки РФ на 2017–2019 гг. (№ 34.3234.2017/ПЧ).

сетью в настоящей работе выступает фрагмент реконструируемой авторами обобщенной терминологической сети, актуальной для терминов категории «Инструмент (искусственный объект)». К рассматриваемой категории относятся термины *квантовая точка*, *нанопленка*, *нановолокно*, *нанопроволока*, *нанопластина*, контексты употребления которых послужили источниками примеров. Основными методами исследования в процессе работы стали контекстный анализ и метод интерпретации полученных данных, которые на начальном этапе способствовали разработке семантических отношений, связывающих термины категории «Инструмент области нанотехнологии» со смежными терминами других категорий. Выделенные семантические отношения между терминами позволили определить характер их системного взаимодействия и далее выстроить некоторые прототипичные модели взаимодействия понятий. Сопоставив набор полученных семантических отношений и прототипичных моделей для вышеперечисленных терминов категории «Инструмент области нанотехнологии», авторам удалось реконструировать фрагмент обобщенной терминологической сети. Данные результаты послужили основанием для формулировки некоторых типовых вопросов, которые следует использовать в качестве поисковых запросов для извлечения научных знаний из экспертной системы. Показано, что сформированные запросы универсальны и позволяют получить необходимую информацию в случае наличия соответствующих семантических отношений в терминологической сети запрашиваемого термина.

Ключевые слова: термин, экспертная система, системное отношение, прототипичная модель, поисковый запрос.

Статья поступила в редакцию 30 мая 2019 г.
Статья рекомендована в печать 23 сентября 2019 г.