# Creating the Applied Subject Area Ontology by Means of the Content of the Digital Semantic Library

## O. M. Ataeva[1*], V. A. Serebryakov[1**], and N. P. Tuchkova[1***]

(Submitted by A. M. Elizarov)

[1]*Federal Research Center "Computer Sciences and Control", Russian Academy of Sciences, Moscow, 119333 Russia*

**Abstract**—This article is devoted to the problem of including a subject area in a semantic library. The subject area is considered, which was not previously presented in part of the ontology of the digital library. A method based on the technology of establishing semantic links with the already accumulated content of the library is proposed. Main idea of the inclusion method is to use dictionaries, thesauri and links to the mathematical encyclopedia. They are used to attach a new scientific publication metadata set to an existing library metadata set. In the course of preliminary processing of texts of scientific publications, their semantic analysis is carried out and a local description of the subject area of this array is formed within the content of the library. Such subject areas as ordinary differential equations, partial differential equations, equations of continuum mechanics, equations of composite mechanics and their solutions expressed in terms of special functions of mathematical physics are considered. The description of the mathematical encyclopedia in its classical form is the central resource. It allows us to identify the appropriate classifiers through the links of the equations of applied problems and their solutions and use them also to form a description of the subject area. This subject area is new and has not been described in our digital library before. Proposed procedure is made possible only by data integration in the library. This takes into account the expansion of the description of subject areas related to the applications of mathematics in interdisciplinary research and technology. The implementation of this procedure is one of the new directions for filling the semantic libraries. It uses the approach of data integration and saturation with connections that are revealed only in the process of analyzing new data. If new references are found in the library during the process of adding a new description, then this method will also make it possible to identify and fill in the gaps in the description of related areas of mathematics, such as applications of solutions to classical problems of mathematical physics.

## 1. INTRODUCTION

There are quite a lot of studies of semantic relations in subject areas and representation of the resulting taxonomies in digital resources in the form of ontologies. Many scientific studies are devoted to data integration to obtain the most voluminous digital representation in the WEB. Mathematics, as a tool for many scientific studies, is especially in demand in the digital environment at all stages of education, from school, university, etc. In the field of mathematical sciences and applications, general education, training, and professional resources are being created. Professional resources are intended for specialists, but like the others, the more popular ones can be used to gain expert knowledge.

*E-mail: oli@ultimeta.ru
**E-mail: serebr@ultimeta.ru
***E-mail: natalia_tuchkova@mail.ru

Among general "educational" resources, Wikipedia[1] is the most popular, and Britannica[2] is the most authoritative, since it provides verified encyclopedic information. In the field of educational resources, numerous special manuals are used, textbooks in the digital libraries of universities, for example, the library of Lomonosov Moscow State University[3], Library for Natural Sciences of RAS[4] and many others. It should be noted that learning resources can include library resources and vice versa, but also be developed separately. In mathematical subject areas, as the most developed resources can be noted Zentralblatt[5], Math-Net.Ru[6], the English version of I. M. Vinogradov[7] [1], Mathematical Encyclopedia of the University of Cambridge[8]. The listed resources are mostly integrated resources. That is, these resources contain dictionaries and thesauri of subject areas, as well as links to the literature of the subject area. The technology for creating such information resources is based on ontological design. An example of the implementation of this approach is a system built on the ontology of mathematical knowledge and presented in [2, 3, 15, 16]. The ontology of the subject area allows to combine various sources and integrate them within the framework of a single content of a digital library. A digital library based on ontological design here means the entire set of semantically related data presented in the content of a digital resource [2–8].

It is proposed here to consider the methodology for forming an ontology of an applied subject area based on the content of a digital semantic library. This approach is implemented based on the integration of subject area data into a single semantic structure, combining the links of dictionaries, thesauri, encyclopedias and real scientific publications of a thematic journal.

The results of this approach are demonstrated by examples of applications of the basic equations of mathematical physics and special functions in the field of problems of composite materials and structures. For example, one of the central problems of the theory of elasticity, the "Lame problem", was chosen, and the semantic links of articles within the content of the digital library were established. This made it possible to carry out one of the main search tasks for describing the applied subject area, namely, to provide the ability to find a publication using all library metadata, including classifiers, keywords, synonyms, and formulas.

The article describes the general approach of ontological modeling of the subject area and its subdomain and discusses an example of implementation by means of a digital library. The work is structured as follows: in Section 2, a description of the subject area is given, in Section 3, modeling of its subdomain based on a general ontology. To do this, Section 3.1 defines the purpose of such modeling. Section 3.2 describes a set of vocabularies based on which the domain subdomain is restricted. Section 3.3 defines a thesaurus extension to describe a subdomain of a subject area. Section 3.4 defines the approach used to incorporate new data into a general ontology based on its extended thesaurus. Section 4 shows an example of including subdomain data based on an extended domain thesaurus.

## 2. ABOUT THE SUBJECT AREA OF MATHEMATICS IN THE CONTEXT OF THE DIGITAL LIBRARY

For several decades of digitalization, the term "ontology" has penetrated into various fields of knowledge. Many researchers, scientists and specialists from different fields of knowledge use, develop or apply ontologies as a mechanism for representing subject areas in information systems. The term "ontology", the origins of which can be found in philosophy, was adopted in the research community of artificial intelligence to formally describe areas of knowledge [9, 10].

As mandatory properties of the description of knowledge in a format that should be defined as an ontology are:

(1) a finite and controlled dictionary of concepts and terms, excluding their ambiguous interpretation, and

---

[1] https://www.wikipedia.org/

[2] https://www.britannica.com/

[3] https://nbmgu.ru/

[4] http://www.benran.ru/

[5] https://zbmath.org/

[6] http://www.mathnet.ru/

[7] https://encyclopediaofmath.org/

[8] https://www.cambridge.org/

(2) a strict hierarchy of relations of subclasses of concepts and terms that describe the knowledge of the subject area [9−13].

One of the features implemented in ontologies is the ability to use inference operations to identify new semantic relationships. This is done thanks to the languages of formal logic underlying ontologies. Therefore, an important element of ontologies is a set of axioms that allow you to extract new knowledge using the mechanisms of reasoning and the use of existing relationships.

At this point in time, a huge number of ontologies have already been developed for various subject areas that cover various aspects of human activity. This paper considers the scientific field "Mathematics" and the process of building and constructing an ontology for one of its applied sections in problems of mathematical physics based on the data sources and library content available from the sources [14, 15].

Note that an important feature of already developed ontologies is the possibility of their reuse to create new ontologies. In this work, research is based on the developed ontologies and thesauri of the subject area, which describe the scientific knowledge of the subject area "Mathematics" at various levels of detail [2, 14−17]. The use of a generalized domain ontology model makes it possible to focus on highlighting such metadata that allow designing specific data structures for various scientific subject areas, identifying common approaches to managing and processing this data. This approach allows you to structure and link various resources, extract from them and contextualize a variety of data, turning them into knowledge [9]. To define the subject area, a general knowledge model is used, within which various data sources can be integrated, and various taxonomies of concepts and terms verified by recognized experts in the scientific field are used. The Mathematical Encyclopedia [16] and the ontology of the scientific subject area, which underlies the construction of the LibMeta semantic library [17] together with the industry classifiers MSC [18] and UDC[9], were used as the basic taxonomy.

## 3. ABOUT THE METHODOLOGY AND PROBLEMS OF ONTOLOGY CONSTRUCTION IN A PARTICULAR CASE

Despite the long history of ontological design, there is still no single standard for building ontologies, but there are various methodologies [22, 23, 31]. Methodologies define different concepts at different levels of conceptualization, but converge in the main sequence of steps necessary to build an ontology: defining the purpose of the ontology, highlighting the main top-level concepts and their properties, and highlighting the relationships between them.

In fact, after defining the goals and objectives of ontology formation, three main processes are distinguished

(1) the collection of data from sources and their ontologization;

(2) the representation of knowledge in the form of taxonomies, and;

(3) the implementation and construction of a formal model suitable for machine use in a specific task [23].

### 3.1. Determination of Goals

The purpose of developing an ontology of applications of the subject area is the need to integrate data describing applications of mathematics and scientific research in the field of these applications, reflected in scientific publications. As has been repeatedly noted [2, 4, 10], there is a certain gap in the representation of knowledge in digital resources and their reflection in bibliographic resources. Examples of the integration of mathematical knowledge and publications are Zentralblatt, and the English version of the Mathematical Encyclopedia, the mathematical digital ecosystem OntoMath [19−21]. And numerous bibliographic resources [2−6, 14] serve as examples of individual databases. This is due, of course, to the fact that the goals of these developments are different. Nevertheless, it is impossible not to note the fact that it is important for the scientific community to have exactly data integration, so that, being within the digital subject area, it would be possible to get acquainted with publications on the chosen topic. This possibility is realizable when creating a common semantic library, where, along with terminological links of the subject area, there are links with publications.

---

[9]https://teacode.com/online/udc/

This is especially important in interdisciplinary studies, which do not always fit into the classical classification, which often makes it difficult to find these works. These areas are modern applications of the classical equations of mathematical physics, which some researchers propose to single out in the section "new applied mathematics". While this section does not exist and its scope is not even defined, the creation of an ontology of the applied area of mathematics is an urgent task of ontological design.

Determining the need to create an ontology of the applied subject area of mathematical physics, first of all, it is possible to determine its purpose—this is the use of a digital library of books, journals, publications and various scientific materials and resources for scientific and educational purposes. On the one hand, these resources are the content of the library, and on the other hand, they are sources of knowledge in this area, which requires classification and categorization. Such an ontology will enrich the data with horizontal and vertical semantic links and define implicit links, for example, between tasks and their applications and between persons associated with these resources. It will also allow you to mutually enrich already linked resources by assigning missing classifier codes or assigning keywords. The solution of this problem within the framework of the semantic library will allow us to come to the construction, in particular, of the thesaurus of the subject area, which will expand and be filled with the growth of information accumulated in the library.

### 3.2. Ontologization

The main data sources can be divided into two broad categories: this is data from journals and scientific publications that demonstrate the development of the subject area over time and the second category is textbooks, monographs, dictionaries and classifiers, which contain the main terminology on which this subject area is based. As the first type of source, we used the publications of the "MMKM" journal over the past 25 years, as well as thesauri, encyclopedia articles, and publications previously accumulated in the LibMeta library. The total number of texts in the corpus is about 10,000. As the second type of sources were used books, textbooks, monographs and terminological dictionaries recognized by experts in the subject area:

- textbook of academicians of the Russian Academy of Sciences A. A. Samarsky and A. N. Tikhonov [24]. About 400 concepts were extracted from this source, which are descriptions of the main problems;

- description of the subject area of continuum mechanics is based on the classification of academician of the Russian Academy of Sciences L. I. Sedov [25]. Sedov's classification is used as a dictionary of basic terms and includes about 1200 terms;

- the subject area of composite mechanics is based on the classification of academician of the Russian Academy of Sciences V. V. Vasiliev [26–28] in English and includes about 2500 terms that were extracted and used as a dictionary of the subject area under consideration;

- glossary of terms from the section "Fullerenes" and related areas [29]. The main emphasis of the dictionary is on the use of terms in Russian-language monographs, educational literature, scientific articles and electronic sources. Contains more 850 terms in Russian;

- glossary of terms from the section "Nanotubes" and related fields [30]. The main emphasis of the dictionary is also on the use of terms in Russian-language monographs, educational literature, scientific articles and electronic sources. Contains over 1000 terms in Russian;

- dictionary of polymer composites based on GOST 32794-2014[10]. Contains about 500 terms that are presented in several languages;

- note that the works of L. I. Sedov and V. V. Vasiliev are devoted to the study of equations of the theory of elasticity, special functions of mathematical physics and applications to specific problems. Glossaries of terms based on them are used to account for historical and modern connections in the selected subject area.

So, on the basis of these sources, the main terms and concepts of the subject area were determined, on the basis of which the thesaurus of the subject area is formed.

---

[10]https://docs.cntd.ru/document/1200113813/

```
<rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax- ns#"
xmlns:lbm="http://libmeta.ru/" xmlns:core="http://libmeta.ru/core">
<lbm:Concept rdf:about="http://libmeta.ru/thesaurus/concept/specfunc_67">
<lbm:thesaurus rdf:resource="http://libmeta.ru/thesaurus/SpecFunctions"/>
<lbm:code/>
<core:url>http://libmeta.ru/concept/show/6784</core:url>
<lbm:descriptor>Ламе функции</lbm:descriptor>
<lbm:comment/>
<lbm:non-descriptor>Lam\'{e} functions</lbm:non-descriptor>
<lbm:properties>
<lbm:property rdf:resource="http://libmeta.ru/thesaurus/attribute/theme_odu">
<lbm:value rdf:resource="http://libmeta.ru/taxon/udc#517.589"/>
</lbm:property>
<lbm:property rdf:resource="http://libmeta.ru/thesaurus/attribute/msc">
<lbm:value rdf:resource="http://libmeta.ru/taxon/msc#33E10"/>
</lbm:property>
</lbm:properties>
<lbm:familyRelation type="http://libmeta.ru/relation/family#related">
<lbm:value rdf:resource="http://libmeta.ru/thesaurus/mathencyclopedia#Ламе_функция"/>
</lbm:familyRelation>
<lbm:familyRelation type="http://libmeta.ru/relation/family#related">
<lbm:value rdf:resource="http://libmeta.ru/thesaurus/mathencyclopedia#Функция"/>
</lbm:familyRelation>
<lbm:familyRelation type="http://libmeta.ru/relation/family#related">
<lbm:value rdf:resource="http://libmeta.ru/thesaurus/mathencyclopedia#Ламе_функция"/>
</lbm:familyRelation>
</lbm:Concept>
</rdf:RDF>
```

**Fig. 1.** Fragment of links the term "Lame" from the semantic library.

### 3.3. Taxonomies

Simultaneously with the construction of an ontology, it becomes necessary to limit it within a specific subject area of science. To do this, a set of concepts is introduced that are used to describe this subject area. The corresponding domain terms are associated with these concepts. Most often, these terms are organized as some kind of taxonomy with support for relationships between them. The structure of this taxonomy can vary in complexity depending on the area being modeled and, if necessary, represent a full-fledged thesaurus with all the richness of relationships. In the future, we will talk about thesauri as a means of organizing concepts (knowledge). Terms presented in this form can be used to process available resources. At the same time, links arise between concepts and resources. The scientific data of the subject area is understood as a set of verified concepts of the scientific subject area and the identified links between them and resources [11−13].

Separately, it is worth mentioning that the thesaurus of the subject area can be either the result of the work of an expert group or built by automated means. Questions of compiling domain thesauri are beyond the scope of this work, as is a discussion of methods for identifying relationships.

Below is an example Fig. 1 of a taxonomy from a dictionary of special functions with relationships. The text is presented in the RDF XML syntax, which reflects the connection of the term with the classifier and the mathematical encyclopedia in the LibMeta library.

### 3.4. Ontology Development

The ontology of the applied subject area of mathematical physics is constructed on the basis of the ontology of the Libmeta semantic library. The ontology of the subject area "Mathematics", including a terminological description based on the "Mathematical Encyclopedia", has already been created earlier. This allows us to use an ontology of information resources to describe publications, persons, tasks, and related formulas to describe data retrieved from new data sources, as well as use a thesaurus to represent domain taxonomies and their relationships.

When defining relationships and concepts of the subject area within the framework of the ontology of the semantic library, three approaches are possible:

- A "top-down" approach where the design of concepts and relationships starts from the most top-level concepts.

- A "bottom-up" approach where the design of concepts and relationships starts from the lowest level data, which are grouped and form more general concepts, etc.

**Table 1.** Descriptor "Lame equation"

| Index position | Content of Index position |
|---|---|
| **D:** | Уравнение Ламе |
| **SYN:** | Lame equation |
| **SYN:** | Lamé equation |
| **SYN:** | Уравнение равновесия в перемещениях |
| **SYN:** | Balance equation in displacements |
| **BT:** | Уравнения теории упругости |
| **BT:** | Equations of elasticity |
| **NT:** | Ламе обобщенная краевая задача для градиентной теории упругости изотропных тел |
| **MATH:** | $(\lambda + \mu)\frac{\partial \Theta}{\partial x} + \mu \Delta u_x = 0$ $(\lambda + \mu)\frac{\partial \Theta}{\partial y} + \mu \Delta u_y = 0$ $(\lambda + \mu)\frac{\partial \Theta}{\partial z} + \mu \Delta u_z = 0$ $\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \mu = G, \Delta$ is Laplace operator |

- And a "combined" approach, when the basic concepts of a particular subject area are formulated, its data is cleaned and partially structured (for example, presented in the form of separate taxonomies) and it is necessary to link disparate data within a given subject area, while enriching it and specializing and refining the included resources.

In our case, we used a "combined" approach. Preliminary work on highlighting the top-level concepts of the subject area and the concepts necessary to describe the structure and relationships of its thesaurus was performed earlier when designing the library. And cleaning tasks were performed as part of preparing data for loading.

## 4. REVEALING THE CONNECTIONS OF PROBLEMS AND SOLUTIONS OF APPLIED PROBLEMS WITH THE CONTENT OF THE SEMANTIC LIBRARY

To solve the problem of expanding the content of the semantic library by including a new subject area, it is necessary to identify the links between the applied area and the concepts of a mathematical encyclopedia. It is also necessary to identify links with already accumulated dictionaries, thesauri, classifiers. The subject area of continuum mechanics and applications to sections of composite materials were considered.

Sources were analyzed and articles of the lexico-semantic index for the equations of the theory of elasticity were compiled. An array of formulas for the selected subject area has been created.

There are two examples of lexico-semantic index for subject thesaurus of "Equations of elasticity" in Table 1 for term "Lame equation" and term "Lame generalized boundary value problem for the gradient theory of elasticity of isotropic bodies" in Table 2.

As an example, we present a small part of the thesaurus devoted to the Lame problem [32–34]. This article is a lexico-semantic index for the subject area of continuum mechanics, section composite materials corresponding to the concept of "Lame generalized boundary value problem for the gradient theory of elasticity of isotropic bodies". The structure of the thesaurus article includes the title, references to synonyms, references to related concepts, references to the literature on which the thesaurus was compiled. The thesaurus entry also includes a set of mathematical formulas related to this concept.

For brevity, we present the values of the positions **"KW"** and **"NOTE"** not in the form of a **Table 2**, but in solid text:

**REF:** *Ильюшин А.А. Механика сплошной среды. М.: Изд-во Моск. ун-та, 1990. 310 с.*

**Table 2.** Descriptor "Lame generalized boundary value problem for the gradient theory of elasticity of isotropic bodies"

| Index position | Content of Index position |
|---|---|
| **D:** | Ламе обобщенная краевая задача для градиентной теории упругости изотропных тел |
| **SYN:** | Lame generalized boundary value problem for the gradient theory of elasticity of isotropic bodies |
| **NT:** | Ламе обобщенная краевая задача для градиентной теории упругости изотропных тел с микроструктурой |
| **NT:** | Lame generalized boundary value problem for the gradient theory of elasticity of isotropic bodies with microstructure |
| **BT:** | Уравнение Ламе |
| **BT:** | Lame equation |
| **MATH:** | Ламе уравнение обобщенное (индексная форма) $L_{ij}(\delta_{jk}R_k - \frac{1}{C}L_{jk}(R_k)) + \rho F_i = 0$, аГаДаЕ $\delta_{jk}$ − дельта Кронекера или $L_{ij}(H_{jk}(R_k)) + \rho F_i = 0$, $H_{jk}(..) = -\frac{1}{C}L_{jk}(..) + \delta_{jk}(..)$, где $H_{jk}(..)$ − оператор Гельмгольца для оператора Ламе (обобщенный оператор Гельмгольца) |
| **MATH:** | Lame equation generalized (index form) $L_{ij}(\delta_{jk}R_k - \frac{1}{C}L_{jk}(R_k)) + \rho F_i = 0$, where $\delta_{jk}$ is Kronecker delta or $L_{ij}(H_{jk}(R_k)) + \rho F_i = 0$, $H_{jk}(..) = -\frac{1}{C}L_{jk}(..) + \delta_{jk}(..)$, where $H_{jk}(..)$ is Helmholtz operator for the Lame operator (generalized Helmholtz operator) |
| **MATH:** | Ламе уравнение обобщенное (векторная форма) $-\frac{1}{C}LL_c(\vec{R}) + \rho\vec{F} = 0$, $L(\vec{R}) = \mu\nabla^2\vec{R} + (\mu + \lambda)\nabla\texttt{div}\vec{R}$, $L_c(\vec{R}) = L(\vec{R}) - C\vec{R}$ |
| **MATH:** | Lame equation generalized (vector form) $-\frac{1}{C}LL_c(\vec{R}) + \rho\vec{F} = 0$, $L(\vec{R}) = \mu\nabla^2\vec{R} + (\mu + \lambda)\nabla\texttt{div}\vec{R}$, $L_c(\vec{R}) = L(\vec{R}) - C\vec{R}$ |

**REF:** *Волков-Богородский Д.Б., Евтушенко Ю.Г., Зубов В.И., Лурье С.А. Численно-аналитический учет масштабных эффектов при расчете деформаций нанокомпозитов с использованием блочного метода мультиполей // Вычислительная математика и математическая физика, 2006, т. 46, № 7, С. 1318−1337.*

**KW:** *Ламе уравнение обобщенное, краевая задача для Ламе уравнения обобщенного, математическая модель межфазного слоя в механике материалов, модель композитных структур с микро- и нано-включениями, модель тонкопленочных наноструктур, представление Нейбера−Папковича обобщенное, Ламе неоднородное уравнение динамики, Ламе однородное уравнение, Ламе оператор, векторная форма Ламе уравнения, индексная форма Ламе*
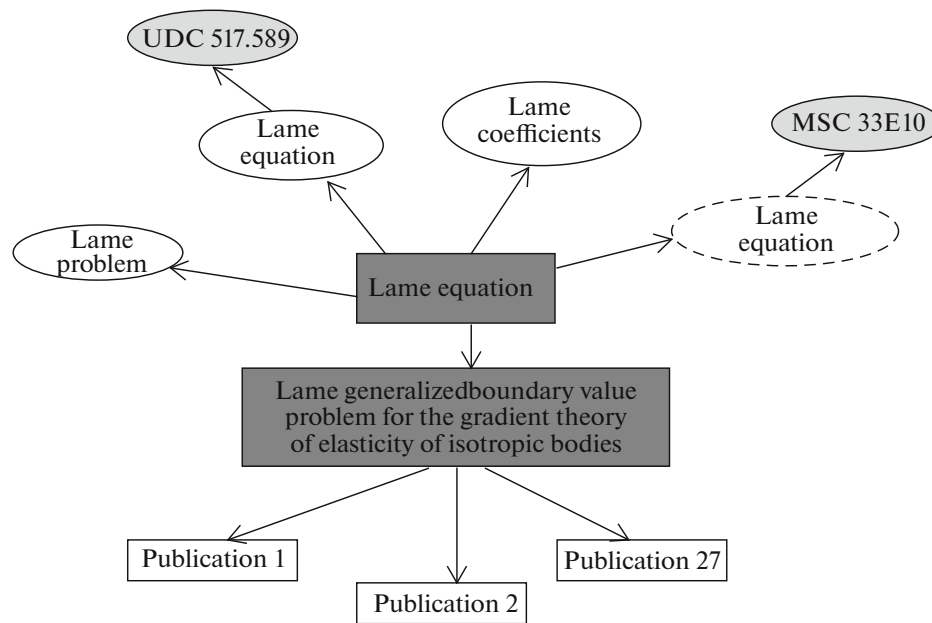
**Fig. 2.** Scheme of the inclusion the publication into subject area of the semantic library.

*уравнения, решение уравнения Ламе представление Галеркина, Ламе волновое уравнение, Ламе полином, Ламе функция.*

**KW:** *Lame equation generalized, boundary value problem for the Lame equation generalized, mathematical model of the interfacial layer in mechanics of materials, model of composite structures with micro- and nano-inclusions, model of thin-film nanostructures, generalized Neuber–Papkovich representation, Lame inhomogeneous equation of dynamics, Lame homogeneous equation, Lame operator, vector form of the Lame equation, index form of the Lame equation, solution of the Lame equation Galerkin representation, Lame wave equation, Lame polynomial, Lame function*

**NOTE:** *Ламе уравнение обобщенное и краевая задача для Ламе уравнения обобщенного определяет математическую модель межфазного слоя в механике материалов или модель композитных структур с микро- и нано-включениями и модель тонкопленочных нано-структур.*

**NOTE:** *The Lame equation generalized and the boundary value problem for the Lame equation generalized determines the mathematical model of the interfacial layer in the mechanics of materials or the model of composite structures with micro- and nano-inclusions and the model of thin-film nanostructures.*

The set of keywords associated with a thesaurus entry is compiled from the associated lookups and dictionaries that were listed in the previous section. These keywords, along with the descriptor and ascriptors of the concept, are used primarily to identify related materials from the content of the library to identify links in text analysis.

Subsets of concepts were identified from the mathematical encyclopedia and the dictionary of special functions for the concept of "Lame equation" first of all (see Fig. 2). These concepts, in turn, are associated with the elements of industry classifiers. They can be used as guidelines for use in classifying publications that do not contain such information.

For example, in this case, the options:

(MSC)-33E10—*Lamé, Mathieu, and spheroidal wave functions*

(UDC)-517.589—*Другие специальные функции и специальные числа*

This list can be extended by allowing the transitivity of relationships and extracting data from the content of the library, which in turn are associated with these elements of the classifiers.

In a new array of data coming to the library (more than a thousand publications), based on the constructed links and thesaurus concepts, related tasks from 27 publications were identified. At the

same time, text analysis, in turn, includes the identification of the main parts of the text in which matches with keywords are most valuable and indicate with high probability the description of tasks of this type in the publication.

Also, by identifying the links between the elements of the thesaurus and the mathematical encyclopedia, in parallel, we get the opportunity to add hierarchical links between concepts to the encyclopedia, based on their links in the thesaurus.

So, on the example of the Lame equation, connections were revealed that made it possible to present the place of this concept in the subject area of mathematics and mechanics within the framework of our data set. Relationships were analyzed using keywords, classifiers and formulas.

In the gray rectangles (see Fig. 2) are the concepts of the constructed thesaurus. In ovals with a solid outline the concept of mathencyclopedia. In ovals with a dotted outline, a concept from the dictionary of special functions. In gray ovals are the codes that we extracted and in the rectangles are the corresponding publications.

## 5. CONCLUSIONS AND FURTHER RESEARCH

The work is devoted to the problem of processing poorly (or insufficiently) structured information, which includes archival articles of national specialized journals. Quite often, they do not contain dedicated sections of annotations, keywords, classifiers and other structural parts of a scientific publication that are standard today. This makes it difficult to find them. However, they are of a certain scientific value as part of the applied scientific field. Modern semantic processing tools make it possible to analyze these texts and supplement them with a digital library in the relevant subject areas. This paper shows how, for one of the applied sections of problems of mathematical physics, the procedure for including arrays of publications of the journal "MMKM" into the ontology of the semantic library based on the available data sources and content of the library is implemented. Connections of a new array of publications with the mathematical encyclopedia, classifiers are obtained, the main concepts of the local subject area of the journal, keywords are identified, which will make it possible to compile a local thesaurus of the journal in the future. Thus, archival articles acquire additional properties, can be presented in search queries and be found thanks to the obtained semantic links.

Now, work on the creation of a semantic library for the considered subject area is ongoing. Further steps will be in the direction of building new links between the main tasks of the subject area and their application. To improve the quality of the description of the subject area in the corpus of texts, it is supposed to process new incoming journal texts to identify their keywords and compare them with the thesaurus of the subject area, built on the basis of the sources described in this article. On the one hand, we get a more accurate description of the subject area, and on the other hand, based on new data, we can offer help to the expert, automating his work in the task of expanding the thesaurus, replenishing it with new concepts and keywords. As a result of the work, we will get a marked up, verified corpus of texts on a given topic, which will allow us to use it in text processing tasks using modern machine learning methods for analyzing mathematical texts and formulas (in Russian aspecially).

## REFERENCES

1. I. M. Vinogradov, *Mathematical Encyclopedy,* Ed. by I. M. Vinogradov (Sov. Entsiklopediya, Moscow, 1979) [in Russian].
2. A. M. Elizarov et al., "Mathematical knowledge ontologies and recommender systems for collections of documents in physics and mathematics," Dokl. Math. **93**, 3231−233 (2016). https://doi.org/10.1134/S1064562416020174
3. A. M. Elizarov et al., "Methods and means for semantic structuring of electronic mathematical documents," Dokl. Math. **90**, 642−645 (2014). https://doi.org/10.1134/S1064562414050275
4. M. R. Kogalovskij and S. I. Parinov, *Semantic Structuring of the Content of Scientific Digital Libraries Based on Ontologies* (NGU, Novosibirsk, 2015) [in Russian].

5. M. Doerr et al., "The European Data Model (EDM)," in *Proceedings of the Torld Library and Information Congress: 76th IFLA General Conference and Assembly* (2010), pp. 10−15.
6. V. Serebryakov, "Cemantic digital libraries. What is it?," CEUR Workshop Proc. **1297**, 1−5 (2014).
7. I. F. Bogdanova and N. F. Bogdanova, "Digital libraries: History and modernity," in *Proceedings of the Joint Scientific Conference Internet and Modern Society* (2017), Vol. 1, pp. 133−154.
8. O. M. Ataeva, "An information model of LibMeta semantic library," Program. Produkty Sist. **4**, 36−44 (2016).
9. T. Gruber, "Ontology of folksonomy: A mash-up of apples and oranges," Int. J. Semantic Web Inform. Syst. **3** (1), 1−11 (2007).
10. T. A. Gavrilova and V. F. Horoshevskij, *Knowledge Bases of Intelligent Systems* (Piter, St.-Peterburg, 2000) [in Russian].
11. M. M. K. Hlava, "The taxobook: History, theories, and concepts of knowledge organization, Part 1 of a 3-part series," Synth. Lect. Inform. Concepts, Retriev., Serv. **6** (3), 1−80 (2014).
12. M. M. K. Hlava, "The Taxobook: Principles and practices of building taxonomies, part 2 of a 3-part series," Synth. Lect. Inform. Concepts, Retriev., Serv. **6** (4), 1−164 (2014).
13. M. M. K. Hlava, "The Taxobook: Applications, implementation, and integration in search: Part 3 of a 3-part series," Synth. Lect. Inform. Concepts, Retriev., Serv. **6** (4), 1−156 (2014).
14. O. M. Ataeva, V. A. Serebryakov, and N. P. Tuchkova, "Mathematical physics branches: Identifying mixed type equations," Lobachevskii J. Math. **40**, 876−886 (2019). https://doi.org/10.1134/S1995080219070047
15. O. Ataeva, V. Serebryakov, and N. Tuchkova, "Ontological approach: Knowledge representation and knowledge extraction," Lobachevskii J. Math. **41**, 1938−1948 (2020). https://doi.org/10.1134/S1995080220100030
16. O. Ataeva, V. Serebryakov, and E. Sinelnikova, "Thesaurus and ontology building for semantic library based on mathematical encyclopedia," in *Proceedings of the DAMDID/RCDL 2019 Conference, Kazan, Russia, October 15−18, 2019,* CEUR Workshop Proc. (2019), pp. 148−157.
17. V. Serebryakov and O. Ataeva,"Ontology based approach to modeling of the subject domain "mathematics" in the digital library," Lobachevskii J. Math. **42**, 1920−1934 (2021). https://doi.org/10.1134/S199508022108028X
18. C. Lange et al., "Reimplementing the mathematics subject classification (MSC) as a linked open dataset," in *Proceedings of the International Conference on Intelligent Computer Mathematics* (2012), pp. 458−462.
19. A. Kirillovich et al., "OntoMathEdu: Towards an educational mathematical ontology," in *Workshop Papers of the 12th Conference on Intelligent Computer Mathematics CICM"2019, Prague, Czech Republic, July 8−12, 2019,* Ed. by C. Kaliszyk et al., CEUR Workshop Proc. **2634** (2020).
20. A. M. Elizarov et al., "Mathematical knowledge representation: Semantic models and formalisms," Lobachevskii J. Math. **35**, 348−354 (2014). https://doi.org/10.1134/S1995080214040143
21. A. M. Elizarov et al., "Digital ecosystem OntoMath: Mathematical knowledge analytics and management," Commun. Comput. Inform. Sci. **706**, 33−46 (2017). https://doi.org/10.1007/978-3-319-57135-5_3
22. M. Bravo, L. F. Hoyos Reyes, and J. A. Reyes Ortiz, "Methodology for ontology design and construction," Account. Administr. **64** (4), 1−24 (2019).
23. D. Allemang, J. Hendler, and F. Gandon, *Semantic web for the Working Ontologist: Effective Modeling for Linked Data, RDFS, and OWL* (ASM, New York, 2020).
24. A. A. Samarskii and A. N. Tikhonov, *Equations of Mathematical Physics,* 7th ed. (Mosk. Gos. Univ., Moscow, 2004; Dover, New York, 2011).
25. L. I. Sedov, *A Course in Continuum Mechanics* (Nauka, Moscow, 1970; Wolters-Noordhoff, Groningen, 1971), Vols. 1, 2.
26. V. V. Vasiliev and Yu. M. Tarnopolskii, *Composite Materials. Reference Book* (Mashinostroenie, Moscow, 1990) [in Russian].
27. V. V. Vasiliev, *Mechanics of Structures Made of Composite Materials* (Mashinostroenie, Moscow, 1988) [in Russian].
28. V. V. Vasiliev and E. V. Morozov, *Advanced Mechanics of Composite Materials and Structures,* 4th ed. (Elsevier, Amsterdam, 2018).
29. E. I. Moiseev, A. A. Muromskii, and N. P. Tuchkova, *Fullerenes in Terms* (MAKS-Press, Moscow, 2010) [in Russian].
30. E. I. Moiseev, A. A. Muromskii, and N. P. Tuchkova, *Nanotubes in Terms* (MAKS-Press, Moscow, 2012) [in Russian].
31. O. Lassila and D. McGuinness, "The role of frame-based representation on the semantic web," Linkoping Electron. Articles Comput. Inform. Sci. **6** (5), 005 (2001).
32. G. Bateman and A. Erdelyi, *Higher Trancsendent Functions* (McGraw-Hill, New York, 1955).
33. M. V. Fedoryuk, "The Lamé wave equation," Russ. Math. Surv. **44**, 153−180 (1989).
34. A. K. Galinysh, "Calculation of plates and shells according to refined theories," Issled. Teor. Plastin Oboloch. **5**, 66−92 (1967).