The Publication of a Knowledge Organization Classification System as a Linked Data Vocabulary.

Olga Lavrenova (0000-0002-4779-9222), Vasily Pavlov (0000-0002-0505-6797)

The Russian State Library, Moscow, Russian Federation.
lavr@rsl.ru, pavlovvv@rsl.ru

Abstract. The report presents an implementation of the Russian Library Bibliographic Classification as a Networked Knowledge Organization System. The project goal is to support the content and indexing; enable information search based on structured knowledge; provide a semantic roadmap for the user; offer the tools to ensure the system's interoperability with other resources of the same nature (Linked Data Vocabularies) in the global network environment.

The project is being implemented by the Russian State Library (RSL) based on its own innovative ideas. The RSL General Classified Catalog was selected as the main source of data for the Classification system. Thus, the Classification System contains not only reference classification numbers from tables, but also numbes created by personnel when cataloging specific publications and manuscripts. To this end, catalog guide cards were fully digitized, approved by experts and embedded into the system. The meaning of each classification number is expressed by a complete string (sequence) of wordings (captions), rather than the last level caption alone. To enrich user queries for search of resources, our library has published its Classification System in the form of Linked Open Data and is now working to enable its smooth integration with other vocabularies.

Keywords: semantic road map; universal library bibliographic classification; enrichment of query; Knowledge Organization System; digital libraries; Linked Open Data

1. Introduction

Let us review the Russian State Library (RSL) [1] information resources, where our Knowledge Organization Classification System supports and will support subject search.

The RSL is the world's largest public library. Any citizen of Russia or any other state who has reached the age of 14 can join the Library as a reader. The RSL houses a unique collection of Russian and foreign documents in 367 languages. The Library comprises over 45,5 m items making it the largest library in Europe in terms of the collection volume. It is also a repository for specialized collections of maps, musical scores, musical recordings, rare books, dissertations, newspapers and other publications. The Library's online catalogue comprised 9 m records. The number of full-text digital documents included in the RSL digital library totals 1,3m titles. The Russian State Library has been a unique repository for dissertation originals on all subjects and in all specialization areas, except for Medicine and Pharmacy, since 1944. The collection contains more than one million volumes. Most of them are included in the Digital library. Currently, a major project aimed at the consolidation of digital documents in the Russian Federation.

The National Digital Library (NDL) [2] is the largest legitimate collection of digitized documents from Russian state libraries, museums and archives. The NDL is a web portal launched as part of the Russian State Programme 'Information Society Development Strategy 2017-2030'. The NDL distributed architecture allows over 100 participants to provide free access to a complete NDL collection at over 4000 libraries in Russia. Web users can only view public domain materials. The NDL catalogue

contains over 39m records allowing locating printed copies in participating public libraries all over the country. Currently, the NDL contains over 4m digitized documents including over 900 thousand dissertations and 2,3m patents. Every month ab to 200 thousand users produce over 700 thousand views.

2. The universal Russian Library Bibliographic Classification as a base for the Subject search

To support subject search in the RSL digital resources, classification numbers of the universal Russian Library Bibliographic Classification [3] for science libraries are used. International library practice traditionally uses acronym "BBK" derived from the Russian acronym "ББК" (Classification Scheme code *rubbk* in MARC 21 format). BBK comes in several versions (complete, medium-level, reduced, etc.). The complete version represents a carefully developed semantic structure. It was not only librarians, but also scientists and specialists working in various areas who joined their efforts to create it. The classification tables are notable for their considerable development depth. Classification numbers generated for specific publications and manuscripts in the course of cataloguing may contain up to 20 and more hierarchy levels. It enables a sufficiently precise coverage of document contents. If the words in LBC chains and controlled vocabularies are not sufficient for a more precise coverage of the document content, bibliographic records are supplemented with non-controlled keywords.

Interesting results have been obtained from the studies to identify lengths classification numbers most frequently used in RSL OPAC conducted through program-based analysis of all the instances where numbers were included in bibliographic records. By way of example, Fig.1 shows data on the amount of BBK numbers of each level (i.e. numbers of a specific length) in the online catalogue for section Ш (SH). "Philological Sciences. Belles-lettres". It presents the diagram of BBK numbers distribution in terms of indexing depth in OPAC records. For section SH, 652361 indices were analyzed, the average indicator was 12.3 levels. The resulting picture, which is quite similar across various knowledge areas, suggests that the most frequently used BBK numbers are those containing 8-12 hierarchy levels.

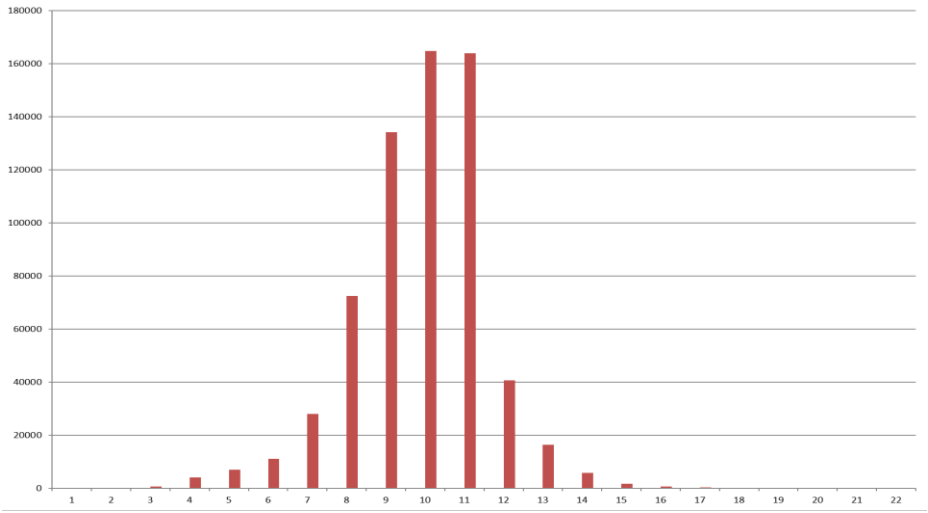

Fig.1. The diagram of BBK numbers distribution in terms of indexing depth. Section SH (Ш).

To support the description and retrieval of universal library information resources in the RSL and other Russian libraries innovative techniques and methods based on hierarchical classification trees was developed [4-6]. The BBK Classification serves as a basis for the enrichment of search queries in the RSL OPAC [7]. At the same time, this technology supplements the search function in the RSL digital library [8]. A full description of a BBK classification number is introduced into a bibliographic record, and all the words therein are used in the query processing.

The majority of modern digital libraries still use document search by words in texts,

keywords and subject headings. The path to the search result in that case looks like a fantastic teleportation to some bibliographic record or a full text document. Many users, however, count on ready-made paths (roads) to data in OPACs or digital libraries. A classification system has to serve as a semantic road map and provide for common orientation of users of traditional and digital libraries in a information retrieval process. This is possible thanks to the classification system's ability to manage semantic relationships between search terms (see the bibliography [9]).

The study of search results in digital libraries by full texts and classification shows that many scientific texts cannot be found, because authors do not introduce into their texts broader (higher-level) terms for those used in the text. They rely on the knowledge the reader already has. It results in a significant loss of information during the search. The main purpose of classifications, thesauri, ontologies is to represent knowledge about the world in systems. Full text search is no exception. Moreover, there are many users globally, who like to create their own paths to the required texts. This is possible only if the system is equipped with a classification hierarchical tree or another knowledge model. This search method cannot be replaced by full-text search. A well-organized interface and data quality make classification systems and full text retrieval mutually complementary.

The RSL has chosen the following 3 technology options [6] to ensure the use of semantic relationships for information retrieval:

A) the inclusion of wordings (captions) of BBK classification numbers in every bibliographic record or metadata embedded in full texts;

B) the creation of a Navigation system for navigating through classification hierarchical trees linked with OPAC;

C) the publication of the Classification System in the form of Linked Open Data.

The technology Option A has been embedded in the RSL OPAC since 1998 year, Option B was developed in 2013-2014 and Option C was implemented this year.

3. Hierarchical strings of captions for classification numbers as document metadata

Option A [4, 6] ensures an enrichment of a bibliographic record with hierarchical relationships fixed inside a classification number caption. The caption of a BBK classification number looks like a captions hierarchy, i.e. a hierarchical sequence or string of captions for all hierarchy levels in a classification number (examples 1 and 2). It enables automatic use of all hierarchical relationships during information retrieval.

As classification number captions are provided directly in the records, hierarchical relations among subjects work automatically (by default) during the search, so the user does not have to think about the way of getting the result. Technologically, it creates the impression that the search is performed simply by the keywords specified. Were it not for hierarchical strings, information losses during subject search that are hidden from the user could be up to 90%, which has been established through search experiments.

Example 1. A classification number with its string of captions from a bibliographic record in OPAC . All the words at all hierarchical levels of the string serve as access points.

Classification number (BBK number): *D236.471* String of captions: *Earth science -- Geophysical sciences (geophisics) --Meteorology -- Atmospheric electricity -- Electric currents in the atmosphere -- Thunderstorms -- Lightning*

The hierarchical string elements are separated by a specified combination of characters: "space" "2 hyphens" "space".

A relevant book can be found in the OPAC, in particular, by the terms: *"Atmospheric electricity" or "Thunderstorms"*.

Example 2 [5]. BBK description of a dissertation.

BBK numbers:

SH141.2-32
SH143.21-32
SH145.3-32
Title:
The "preposition" category in the Russian, English and Modern Greek. Dissertation.
Strings of captions:
1) Philological sciences. Belles-lettres -- Linguistics – Languages of the world -- Indo-European languages -- Slavic languages – Eastern slavic languages -- Russian -- Lexicology – Semantics (semasiology)
2) Philological sciences. Belles-lettres -- Linguistics – Languages of the world -- Indo-European languages -- Germanic languages – English -- Lexicology – Semantics (semasiology)
3) Philological sciences. Belles-lettres -- Linguistics – Languages of the world -- Indo-European languages -- Greek -- Modern Greek -- Lexicology – Semantics (semasiology)
Keyword*:*
preposition

The dissertation can be found in the RSL digital library, in particular, by the terms: *"lexicology" AND "Greek"or "semasiology" AND "germanic" AND "slavic" AND "languages"*.
Of course, such possibilities are created only by hierarchical caption strings.

4. Navigation System

The RSL General Classified Catalogue (GCC) was selected as the main data source for the classification system. It is a conceptual basis of our knowledge system. This approach fundamentally differentiates the RSL's model from conventional implementations of classification systems for search support in online catalogues. Our system partly simulates the subject search procedure used in traditional classified catalogues. Records are found there under the selected "guide card", i.e. on the selected hierarchy level only as a result of search. The user enters a set of words, the system finds all the strings containing such words in any grammatical forms and displays the subjects (strings) on the screen together with their classification numbers. The user selects the required subjects, obtains the data on the number of documents available on the subjects concerned and decides which of them should be sent for search in the online catalogue. Furthermore, the user can navigate up and down from the hierarchy level found. In response, descriptions of the documents available in the RSL in traditional and electronic formats are displayed.

Classified catalogue guide cards (totaling about 130,000 units) were fully digitized, checked, edited, modernized by experts and embedded into the Navigation system (Navigator).The data concerned were normalized for the use in bibliographic records. Thus, a kind of controlled vocabulary was created. Furthermore, the RSL staff may add new useful structural elements from the BBK standard to BBK numbers and caption strings as part of indexing processes without changing the elements already normalized. New strings can be included in the normalized list by experts.

As is well known, some libraries implement various machine-readable classification schemes in their initial forms [10, 11]. However, users may find it difficult to design search characteristics using the algorithms built into a catalog, if the classification system has a complex structure. We use classification numbers created by professionals.

It is known that one of the key features of the model of any object is its adequacy for a specific task. Where another task needs to be performed, a different model will be required to reflect the object properties. From our perspective, knowledge models, such as machine-readable standard classification tables and hierarchical trees of virtual classified catalogue guide cards have different purposes. While the first model enables the generation of classification numbers during the processing of documents, the second one provides ready-made paths to the documents contained in data bases using ready-made classification numbers available for specific documents. The purpose of the model presented in this report is to support text search by classification numbers

which have already been created during the processing of document contents.

The Navigation System supports direct hierarchical search, search by arbitrary combinations of words in various grammatical forms of the Russian language taking into account certain word formation elements.

When processing a query, the Navigator displays [12]:

- classification numbers generated in bibliographic records during the cataloguing process;
- the caption hierarchy (string) for each classification number;
- classification hierarchical tree levels in consecutive order;
- the number of relevant bibliographic records in the RSL OPAC for each classification number found.

Since the Navigator technologies are transferred to the LOD environment, it is suggested to consider the search examples in the relevant part of the paper. The Navigator transfers the classification numbers found to the catalogue to search for bibliographic records and full-text documents.

Unfortunately, this technology cannot be applied in respect of bibliographic records which for various reasons do not have BBK numbers. Their numbers will be restored in the future.

Where it is questioned whether the transformation of traditional library classifications into modern knowledge models is with the enormous effort put into it, there are at least two arguments in favour of that work:

- we cannot afford to lose the searching tools for millions of documents related to the country's cultural heritage, the content of which is described in the libraries using such classifications; it may lead to the loss of road maps to many publications and manuscripts in the libraries' collections during the subject search;
- the development of new structures for the presentation of large arrays of resources, universal in their content, will require huge expenses and participation of a wide range of specialists; over more than half-a-century-long history of information technologies development, nobody has managed to do it, so far.

5. Classification System in LOD Environment

This LOD technologies development project is supported by the Russian Foundation for Basic Research (RFBR).

5.1. The structure features of classification data

There are three main differences between our classification model and other classification models that exist in LOD:

1) As already pointed out earlier our classification system is not formed on the basis of a certain classification standard with all its components and rules for building BBK numbers by library staff. It is Their places in the semantic space have already been identified by librarians. Naturally, the main classification divisions from the standard tables are also presented in the hierarchy.

2) In RDF representations, a classification number should be linked not only to the hierarchy lower level caption, as it is done in other similar projects [10, 11], but to the complete string of captions for such classification number.

Let us compare a classification with a thesaurus. A thesaurus fixes only paradigmatic relations among its elements [13]. A completely different picture is observed in the hierarchical structure of a classification system. Classification numbers encode semantic elements of varying degrees of complexity including both paradigmatic and syntagmatic links. An analogy can be drawn with natural language sentences, where the semantic content is not equal to the simple sum of the meanings comprising it.

Let us consider number *SH145.3-32* from Example 2.

«Philological sciences. Belles-lettres – Linguistics» - there is a paradigmatic relationship between the elements.

«Languages of the world -- Indo-European languages -- Greek -- Modern Greek» - there are paradigmatic relationships here as well. «*Lexicology – Semantics*» - paradigmatic relationship.

However, there are also syntagmatic links among these fragments which form the semantics of the number. The aforesaid brings us to an important conclusion well known to classification experts: the semantic content of a classification number is expressed by a complete hierarchy of its captions.

3) Only those elements of the BBK classification are represented in our system as RDF triplets that can be used to enrich user queries.

Therefore, our system does not use, for example, comments for the library staff concerning the specificities of constructing numbers, which are found in other classification models. For some notes, as much as a special namespace with several new links was created under the UDC/LOD project [10]. The requirement was to identify such elements of classification data that would improve search quality, provided that they could be attached to the initial query using software, as it is implemented in LOD environment.

As part of the project, additional ways of enriching user queries are being developed. In particular, search words from guidance notes thereto taken from BBK tables are introduced into RDF number representations. They are linked to the concept by tag skos:example.

Example 3. For number Ищ(*SHCH)314.043 (Art. Art history -- Music – Individual kinds of music and music performance – Vocal music – Theory of Vocal Music – Types, genres and forms of vocal music – Vocal genres of chamber music repertoire*) the following guidance note is encoded: "*Cantata, vocal cycles, romance song, song, ballad, etc."*

In the Navigator, information requests were already enriched with various grammatical forms of the Russian words. To download grammatical forms, Russian freely available linguistic databases were used [14, 15].

For comparison purposes, we present some examples not only from the Russian language, but from German and English languages as well.

The Russian language grammar is quite complicated. Different endings are used to build various case forms of various declension paradigms, as well as plural forms.

Example 4. The noun "sun" in various Russian cases.

солнц<u>е</u> - солнц<u>а</u> – солнц<u>ем</u> - солнц<u>у</u> / solnc<u>e</u> – solnc<u>a</u> – solncem – solncu.

There are also suppletive word forms, i.e. the forms of one and the same word formed from different roots.

Example 5. Suppletive plural forms.

человек / chelovek – люди / lyudi (Russian), man – people (English), Mensch – Menschen, Leute (German)

Various word forms are equated to each other during the search in the System.

In addition, word formations of all the words contained on the Classification System database were compared specially for the project.

Example 6. Adjectives formed from nouns.

Солнце/solnce = солнечный/solnechnyj (Russian), Sonne = sonnig (German), sun = sunny (English).

Problems of synonimy for the words contained in the BBK captions is planned to be solved next year.

5.2. The RDF presentation of the classification data

The structure of the RDF presentation for the Classification system looks as follows [12]:

URI - skos:Concept ,
classification number - skos:notation ,
hierarchical string of captions (caption) - skos:prefLabel ,
alternative hierarchical string of captions - skos:altLabel ,
broader call number- skos:broader,
narrower call number - skos:narrower (automatically generated),
see also - skos:related ,
note with examples - skos:example ,

end of caption - skos:hiddenLabel (automatically generated),
service information:
skos:historyNote ,
skos:changeNote.

Example 4. RDF statement for a concept [16]

It is understood that original RDF statements contain Russian words written in Cyrillic. They were translated into English for this publication.

URI for E472.311.5 is http://lod.rsl.ru/bbkgsk/concepts/%D0%95472.311.5.

URI for E472.311 is http://lod.rsl.ru/bbkgsk/concepts/%D0%95472.311.

```
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
 <skos:Concept rdf:about="http://lod.rsl.ru/bbkgsk/concepts/%D0%95472.311.5">
<skos:notation > E472.311.5</skos:notation>
<skos:prefLabel xml:lang="ru"> Life sciencies -- Microbiology -- Physiology biophysics and biochemistry of мicroorganisms -- Biochemistry of мicroorganisms -- Metabolism and energy exchange of мicroorganisms. Nutrition of microorganisms -- Anabolism (assimilation) -- Biosynthesis -- Photosynthesis. Photosynthetic microorganisms
</skos:prefLabel>
<skos: broader xml:lang="ru">
http://lod.rsl.ru/bbkgsk/concepts/%D0%95472.311</skos: broader>
<skos:hiddenLabel xml:lang="ru"> Photosynthesis. Photosynthetic microorganisms
</skos:hiddenLabel>
</skos:Concept>
```

The classification files were prepared by converting data from Navigator into RDF. The files were loaded into the semantic storage for subsequent data manipulation using the SPARQL query language.

The classification system is published under the license of the RSL, based on the Creative Commons recommendations, option CC BY-NC-ND «Attribution-NonCommercial-NoDerivs» [17].

5.3. The software and information retrieval technology

The software includes a software package meeting all the requirements. It is an Apache Jena package that serves as a platform for creating applications for linked data and the Semantic Web (for more details see [18]). To implement the database scheme and enable full text search by BBK number captions, taking into account the morphology of the Russian language, it was decided to migrate to Apache Jena framework.

Apache Jena is an open source project and a semantic web framework. The use of Apache Jena helped us to minimize development costs. The framework offers:

- triple store (TDB);
- SPARQL endpoint (Fuseki);
- full text query support (ARQ).

Apache Jena Fuseki is a SPARQL server. It can run as a operating system service, as a Java web application (WAR file), and as a standalone server. It provides security (using Apache Shiro) and has a user interface for server monitoring and administration. It provides the SPARQL 1.1 protocols for query and update as well as the SPARQL Graph Store protocol.

Fuseki is tightly integrated with TDB to provide a robust, transactional persistent storage layer, and incorporates Jena text query and Jena spatial query. It can be used to provide the protocol engine for other RDF query and storage systems.TDB is a component of Jena for RDF storage and query. It support the full range of Jena APIs. TDB can be used as a high performance RDF store on a single machine. This documentation describes the latest version, unless otherwise noted. A TDB store can be accessed and managed with the provided command line scripts and via the Jena API. When accessed using transactions a TDB dataset is protected against corruption, unexpected process terminations and system crashes.

ARQ is a query engine for Jena that supports the SPARQL RDF Query language.

SPARQL is the query language developed by the W3C RDF Data Access Working Group.

ARQ Features:

- Standard SPARQL
- Free text search via Lucene
- SPARQL/Update
- Access and extension of the SPARQL algebra
- Support for custom filter functions
- Property functions for custom processing of semantic relationships
- Aggregation, GROUP BY and assignment as SPARQL extensions
- Support for federated query
- Support for extension to other storage systems
- Client-support for remote access to any SPARQL endpoint

The use of Apache Jena framework enabled us to resolve in a very short time the following issues: loading, storing and organizing access to project data. Ample features of the framework's application program interfaces (API) allowed us to quickly address the task of putting together a website for the publication of BBK/GCC data, enabling navigation and search, and posting general information on the project. The project website consists of three parts:

- the first part represents an application written in JavaScript implementing the navigator designed to work across the BBK numbers tree;
- the second part is an application written in JavaScript enabling the user to find the required BBK numbers by their captions, keywords;
- the third part comprises a set of static pages featuring project information.

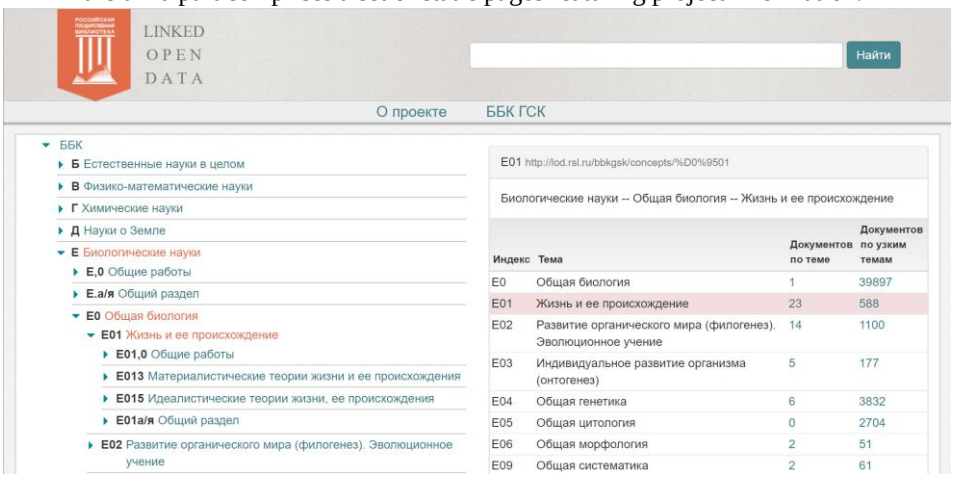

Fig. 2 – BBK Numbers navigation page. *E01 "Life sciences -- General biology - The origin of life*"

The user can navigate through the classification tree on the left side of the screen and select the required BBK number. To the right, he sees a complete string of captions. Below is the amount of bibliographic records for this BBK number, other numbers at the same level and at the higher one.

Example 5. The result of a word search in captions of the Classification system as by keyword - "*Greek*". The user sees on the screen 116 different BBK numbers with caption strings from different sections. The following are some of them:

Physics and Mathematics -- Mathematics -- General Section -- Persons of mathematicians -- Foreign countries - Pythagoras from Samos (580? -500? BC) (Ancient Greece)

Life Sciences -- Paleontology -- Paleozoology -- Paleozoogeography -- Europe -- Southern Europe -- Greece

History. Historical Sciences -- History of Historical Science -- The General History of Historical Science -- The Development of Historical Thought and the Accumulation of Historical Knowledge in the Ancient World -- The Development of Historical Thought and the Accumulation of Historical Knowledge in the Ancient World - The Development of Historical Thought in Ancient Greece

History. Historical Sciences -- Source Studies. Auxiliary Historical Disciplines -- Numismatics - Ancient World -- Numismatics of the Ancient World -- Numismatics of Ancient Greece

Art. Art Studies -- Decorative and Applied Art -- Types of art products that differ in material -- Ceramics -- History of ceramics -- The general history of ceramics--- Ceramics of the Ancient World-- Antique ceramics - Ancient Greece

Art. Art history - Graphics -- Types and genres of graphics -- Engraving -- Types of engravings that differ in material and technique -- Wood engraving (woodcut) -- History of woodcutting -- Selected foreign countries -- Europe - Greece

Art. Arts -- Theater -- Specific theater -- Musical theater- - Opera theater -- History of opera house -- Selected foreign countries -- Europe- - Greece

Religion. Atheism -- Religion -- Individual religions -- Christianity -- Orthodoxy -- Hellas (Greek) Orthodox Church.

The user can also display the amount of bibliographic records or go to the appropriate level of the hierarchical tree.

The both applications written in JavaScript use Apache Jena for BBK numbers search and visualization. Moreover, these applications are integrated with the RSL OPAC from which they receive information about the number of documents under respective BBK numbers. As a result, the user can view the documents on the required subject. The following is a general diagram showing how the user works with the website.

To address the tasks of translating BBK numbers and linking BBK numbers with UDC numbers, a specialized software was developed. As a result, a web application using Django framework was implemented. The resulting web application consists of four modules.

The first module is designed for data collection and import. First of all, the following BBK numbers information was converted and loaded: URI, BBK numbers and their captions.

As the next stage, all the captions were automatically translated using machine translation systems. Yandex, Google and Systran open programming interfaces (APIs) were used as translation systems. Also, for linking with UDC numbers, 1800 upper level classes were downloaded from UDC Summary Linked Data website [10]. These data were also converted and corresponding information on BBK numbers was loaded into the software, except for the captions which were simultaneously loaded in two languages: Russian and English.

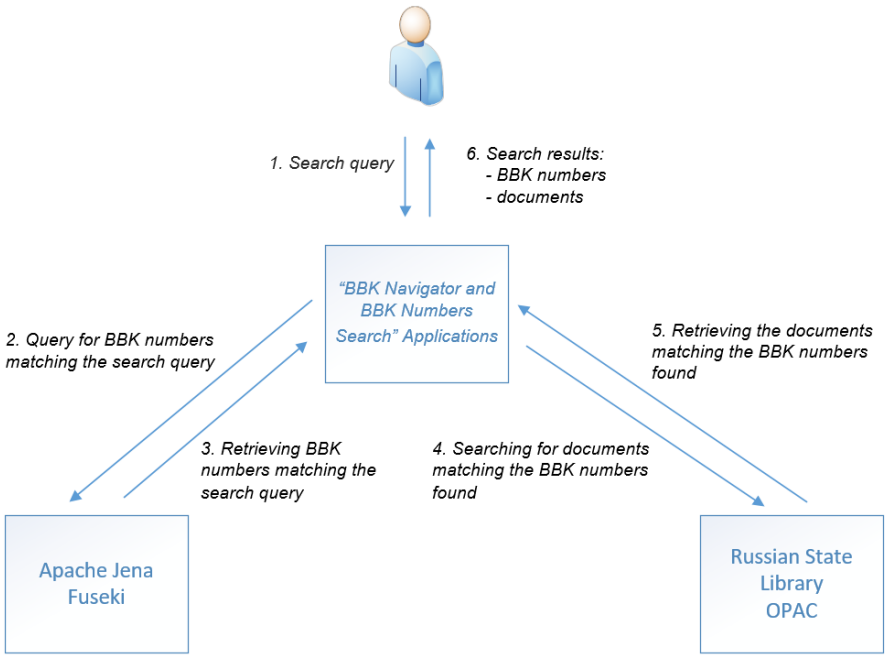

Fig. 3 – Workflow diagram of "BBK/GCC Navigator and Search" Application

The second module represents a database, where all imported data were saved. Also, room for storing the following new data was prepared:

- final translations of the captions;
- data on number links.

The third module represents a web interface for the database. It provides for the visualization of data stored in the second module and enables translation and data linking operations. The module appearance is presented below. Figure 4 shows its

appearance in the translation mode. In the upper part of the application, the operation mode can be selected: either BBK translation or linking of BBK and UDC numbers.

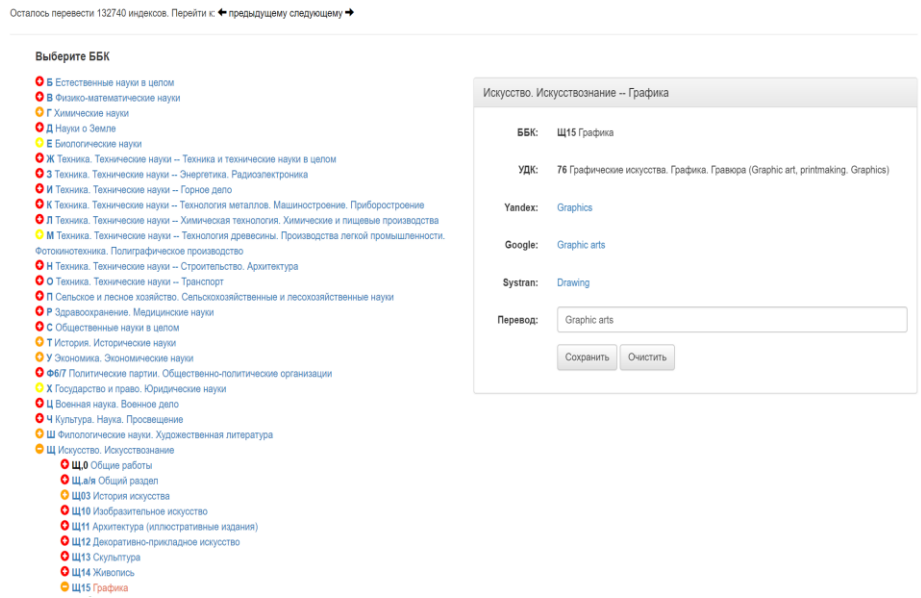

Fig. 4 – Web application page in the BBK numbers translating mode

The BBK numbers tree is located on the left hand side of the central part of the application. The user can select any of the BBK numbers for translation, and if a number has not yet been translated, the right side of the screen will display information on the BBK number and the UDC number linked to it, if such link has been established, as well as the automatic translations made. To translate a number, the user can choose one of the suggested translations and correct it, if the automatic translation suggested was not accurate. Above the BBK numbers tree, there is a button showing how many numbers are left to be translated and allowing the user to proceed to the first number which has not yet been translated.

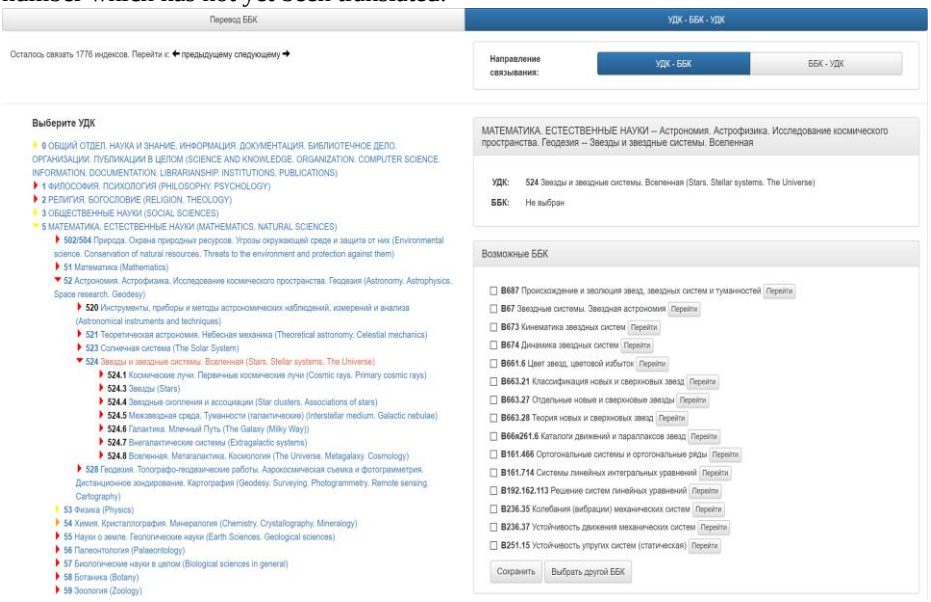

Fig. 5 – Web application page in the BBK numbers linking mode

The functions of the fourth and the last module include data conversion and export to Apache Jena.

The BBK numbers linking functionality was implemented by stages. First, the possibility of linking one BBK number to one UDC number was implemented. However, since the UDC numbers available in open access make about 2 percent of the existing BBK numbers, linking in the UDC-BBK direction is more practical. Therefore, the feature enabling the selection of the linking direction was added. Figure 5 shows a screenshot of the system described with the direction selection feature in the upper right hand side of the screen. The left hand side of the screen, depending on the linking direction, displays either the BBK or UDC tree. The right hand side displays the information on the current number, linked number, if the link has already been

established. Displayed below us information on the automatically selected most matching numbers. The user has the following options:

- he can select the most matching number out of the ones suggested;
- he can go to in the index tree to any of the numbers suggested and detail the number;
- he can go straight to the numbers tree root and select the required number manually.

5.4. Practical Uses of BBK/GCC

Let's consider possible variants of using the BBK/GCC in electronic catalogs of libraries. Below are the options written in increasing complexity:

- Organizing subject search through visualization of the BBK/GCC tree and enabling the reader to navigate through the tree. The easiest option is to display the BBK/ GCC tree for the reader to view, so that he could choose the subjects he is interested in himself. This option can be found in the RSL Linked Open Data website [16] in the BBK/GCC section. The BBK headings tree is presented to the user, and he can find the required subject by opening the branches of the tree, i.e. going from the general subject to more specific topics.
- Displaying subject headings in search tips for subsequent search by BBK numbers. Displaying subject headings as search tips can be visually organized in different ways. For example, it can be done as a pop-up tip under the search box, where the information displayed changes as the user types his search query. Or else, it can be a separate information block which will appear on the search results page.
- Displaying the literature list by keywords. Another option is the direct document rather than BBK number search. In that case, the user's query is processed, the most matching BBK numbers are found and used for subsequent search in any online catalogue. As a result of such search, the user gets the documents on the required subject.

6. Conclusion

We can now answer the question whether the traditional knowledge organization system can improve full text retrieval? It certainly can.

The enrichment of bibliographic records and full text digital resources supplemented by keywords and semantic relationships corresponding to the adopted science structure ensures search completeness and the availability of new important information services.

We are considering the application of the Classification system in LOD environment in the following areas:

- enriching queries from various users (organizations and individuals) to enable their transfer RSL OPAC and digital library;
- enriching queries from various users to enable their transfer to OPACs of other Russian libraries using the complete version of BBK tables;
- presentating the medium-level version of BBK tables in RDF; providing links between two table versions through upper-level classification indices and words for enriching queries submitted to other types of Russian libraries;
- enriching queries with UDC (Universal Decimal Classification) [10] and DDC (the Dewey Decimal Classification) indices [11], as many Russian and foreign libraries are using them (see also [19]);
- enriching queries containing UDC indices from the tables published in the LOD environment using BBK numbers to perform searches in the RSL OPAC;
- enriching queries submitted to Yandex search engine with words from our Classification system associated with the words used in the query (in Russia, Yandex is used on a par with Google for Internet searches);
- addition of Wikipedia articles with links to publications in the RSL and,

conversely, the opportunity for the RSL user to provide articles from Wikipedia;

- using the Classification system for searches in the National Digital Library.

It is essential for the development of LOD technologies to solve the problem of translating our hierarchical captions into English. It is required in order to enable the processing of queries from foreign users of our online catalogues and digital libraries in the LOD environment. Moreover, such data will be used for computer comparisons of heading names of various classifications to form links among classification numbers.

We naturally need discussions of practical initiatives aimed at forming links among congruent vocabularies in other libraries. Information on multilingual and interdisciplinary KOS applications and tools is also very important for our project.

References

1. Russian State Library. Official websit. http://www.rsl.ru
2. National Digital Library (NDL). http://нэб.рф
3. Library-Bibliographical Classification (LBC)/ The national classification system of the Russian Federation. http://www.iskoi.org/doc/lbc.htm
4. Lavrenova O.A. (2004). Subject Search in OPACs and Digital Libraries. Bibliotekovedenije (Library and information science). ISSN 0869-608X. 5:42-50
5. Lavrenova O.A. (2012) The National Library Bibliographic Classification (BBK) as a Basis for Subject Search in the Integrated RSL Digital Library. Project Presentation. Networked Knowledge Organization Systems and Services. The 11th European Networked Knowledge Organization Systems (NKOS) Workshop. TPDL Workshope, Paphos, Cyprus. https://www.comp.glam.ac.uk/pages/research/hypermedia/nkos/nkos2012/programme2012.html
6. Lavrenova O.A. (2013). User's Options for Information Retrieval in Digital Libraries or "a Knight at the Crossroads". Bibliotekovedenije (Library and information science). ISSN 0869-608X. 3:43- 52
7. Russian State Library. OPAC: http://aleph.rsl.ru/
8. Russian State Library. Digital Library: http://search.rsl.ru
9. OPAC semantici. Bibliografia a cura di Emanuela Casson, Claudio Gnoli, Riccardo Ridi, Aida Slavić, Giulia Visintin aggiornamento 2015.06.30 vedere anche le pubblicazioni prodotte dal progetto Opac semantici. http://www-dimat.unipv.it/biblio/sem/bibl.htm
10. UDC Summary Linked Data: http://udcdata.info
11. Dewey Decimal Classification (DDC): https://datahub.io/dataset/dewey_decimal_classification
12. Lavrenova O.A. and Pavlov V.V. (2017) Library Bibliographic Classification as a Traditional Knowledge Organization System in Linked Open Data Environment. Naucnije i tekhnicheskije bibliotiki (Scientific and Technical Libraries). ISSN 0130-8765. 4: 44-59
13. RDF Schema 1.1. W3C Recommendation 25 February 2014 - httpshttps://www.w3.org/TR/rdf-schema/Advantages of thesaurus representation using the Simple Knowledge Organization System (SKOS) compared with proposed alternatives . Juan-Antonio Pastor-Sanchez, Francisco Javier Martínez Mendez and José Vicente Rodríguez-Muñoz. http://www.informationr.net/ir/14-4/paper422.html
14. Catalog of Russian linguistic resources NLPub. https://nlpub.ru
15. Ustalov D. A. Russian Thesauri as Linked Open Data. http://www.dialog-21.ru/media/1103/ustalovda.pdf
16. BBK Linked Open Data: http://lod.rsl.ru/
17. Creative Commons. https://creativecommons.org/licenses/?lang=ru
18. Apache Jena. https://jena.apache.org/
19. Mapping Classification Schemes, D.Hiom, DESIRE report. http://www.desire.org/results/discovery/cat/mapclass_des.htm