Ontology-related Complex for Semantic Processing of Scientific Data

Oleksandr Palagin[1], Mykola Petrenko[1], Mykola Boyko[1]

[1] *Glushkov Institute of Cybernetics of National Academy of Sciences of Ukraine, 40 Glushkov ave., Kyiv, 03187, Ukraine*

Abstract

Developing theories, methods and algorithms to discover and generate new knowledge has always been one of the most important tasks of any researcher, especially if they are actively working on the creation of new scientific publications. There is no universal language for the formal description of concepts (knowledge) and systemology of transdisciplinary scientific research. This raises a set of topical problems for researchers, and one of them is the way to speed up information (in the form of cognitive-structure) search process in their own sources. The ontology-related complex for semantic processing of scientific data is designed to solve this problem for the researcher who has dozens to hundreds of published scientific papers. We are not aware of any search engines that could provide the same information for a researcher in such a short time. The ontology-related complex implements information retrieval and knowledge discovery technologies with emphasis on technologies and tools such as Semantic Web and cognitive graphics. The development of such a complex consists of three stages. The first stage creates the instruments for complex development, methods and algorithms for interaction between the components of the system "User — Knowledge engineer — Remote endpoint", also at this stage data is added to the system. The second stage solves the task of multimedia representation for conceptual and figurative structures described in scientific documents. The third stage solves the task of acquiring new knowledge.

Keywords

Transdisciplinary scientific research, Semantic Web technology, ontological engineering, database of scientific publications.

1. Introduction

There are many applications for searching information from various databases (DBs), including specialized ones. Most of these applications do not take into account the cognitive aspect of data processing, which is necessary for creativity, in particular for the researcher (RSR).

A separate problem is the multimedia (conceptual and figurative) representation of search results and their comparison with the conceptual structure of the Knowledge Domain (KD); this is of interest in order to gain new knowledge. The processing of scientific publications by a single author, authors of a scientific unit or institute using Semantic Web technology is relevant for scientific research.

The ontology related complex (OrC) for semantic processing of scientific data uses information retrieval and knowledge discovery technologies with a focus on Semantic Web and cognitive graphics technologies and tools [1–3]. These technologies and related tools enable the creation of multimedia representations of conceptual and figurative structures that are described in scientific articles. We use scientific publications and databases containing them (DBSP) as scientific data. Semantic Web technologies allow the creation and processing of a Resource Description Framework (RDF) repository of scientific publications, the creation of local or remote endpoints, and execution of SPARQL-queries. Of the plethora of Semantic Web technologies, it is necessary to highlight the SPARQL-technology, which allows the RSR to create queries of arbitrary complexity and receive information in response.

[1] 13th International Scientific and Practical Conference from Programming UkrPROGP'2022, October 11-12, 2022, Kyiv, Ukraine
EMAIL: palagin_a@ukr.net (A. 1); petrng@ukr.net. (A. 2); swsfrmac@gmail.com.; (A. 3)
ORCID: 0000-0003-3223-1391 (A. 1); 0000-0001-6440-0706 (A. 2); 0000-0003-1723-5765 (A. 3)

A summary chart for the development of the OrC DBSP is shown in Figure 1. It includes a preparation stage block and main stage blocks with variations A, B, and C. The preparation stage is described in detail in [1]. The ontological graphs of subject area and one of scientific publications that serve as data for implementation of the main stage, variation B phase 2 are also given there. "New knowledge" in the phase 3 description is a knowledge that is not contained in the DBSP of the user (author).

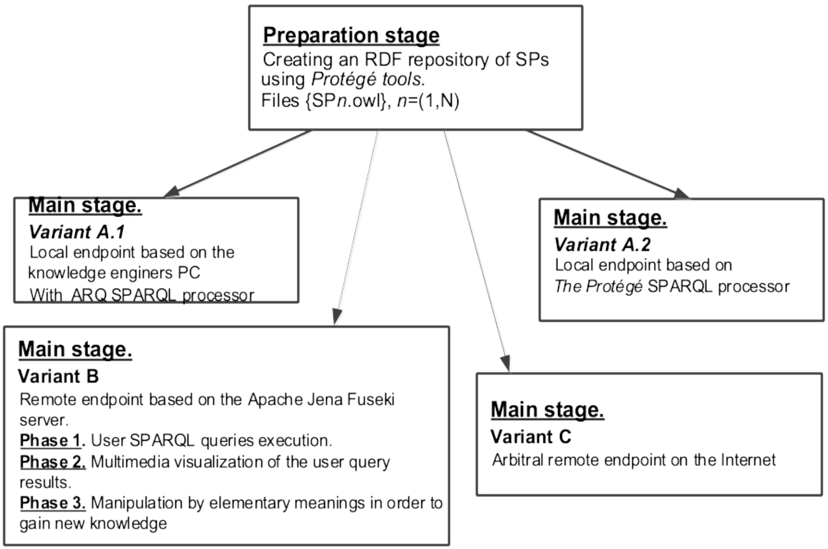

Preparation stage
Creating an RDF repository of SPs using Protégé tools.
Files {SPn.owl}, n=(1,N)
Main stage.
Variant A.1
Local endpoint based on the knowledge enginers PC
With ARQ SPARQL processor
Main stage.
Variant A.2
Local endpoint based on
The Protégé SPARQL processor
Main stage.
Variant B
Remote endpoint based on the Apache Jena Fuseki server.
Phase 1. User SPARQL queries execution.
Phase 2. Multimedia visualization of the user query results.
Phase 3. Manipulation by elementary meanings in order to gain new knowledge
Main stage.
Variant C
Arbitral remote endpoint on the Internet

Figure 1: Summary chart for the development of the OrC DBSP

We are aware of personalized knowledge base of the researcher, in which a number of functionalities declared, these features support the processes of scientific and creative activity [2]. Such personalized knowledge base is:

- a tool to support scientific research, and one of the central areas of development of practical informatics [4, 5];
- development of a knowledge system for the RSR, in order to create new knowledge (or to streamline existing knowledge, check errors and inconsistencies etc.) [6–9];
- one of the main subsystems of a modern research design system [10], an automated workstation for the researcher [4];
- one of the basic elements for establishing a permanent canonical knowledge [11] and supporting the functioning of a knowledge-oriented information system [12].

There is a close relationship between Semantic Web and UML technologies. In particular, we are talking about the relationship between the Web Ontology Language (OWL) syntax and the visual modeling of Unified Modeling Language (UML) diagrams. UML is presented as a general-purpose visual modeling language, which is designed to specify, visualize, design and document components of software, business processes and other systems. UML is a simple and powerful modelling tool, that can be used effectively to create conceptual, logical and graphical models of complex systems that are built for different purposes. This language incorporates all the best practices and qualities of software engineering that have been successfully used for many years, to model large and complex systems [13].

Visual modeling in UML can be represented as a process of gradual descent from the most general and abstract conceptual model of the original system to the logical and then to the physical model of the corresponding software system. For these purposes, first a model is built in the form of a so-called use case diagram. This diagram describes the functional purpose of the system, what the system will do in the course of its operation. A use case diagram is an initial conceptual representation, or conceptual model, of a system during its design and development [14, 15].

OrC "Database of scientific publications" was created for the author, who is actively involved in the preparation and release of new SPs. Of course, it is possible to search in your own SPs manually (which happens in most cases), but with the help of the OrC this search can be greatly accelerated. In addition, it is possible to automatically structure the retrieved data into appropriate templates for future SPs.

We will now discuss the development of architectural, structural components and UML-diagrams. Diagrams showing the operation of the OrC based on the Apache Jena Fuseki remote endpoint. In addition, we will discuss examples of how to use a formally described scientific paper, and perform a number of queries to some papers.

The purpose of this paper is to show the process of the OrC development. The complex allows us to greatly accelerate the search for information by the user (in his own DB of SPs), gives visual representation of the publication concepts and the relevant subject area, and implements Brooks' famous formula for gaining new knowledge [7, 8]:

$$K(S) + dI = K(S + dS),$$

where $K(S)$ is the original knowledge structure, which is modified by the results of processing of the information portion dI, creating a new structure $K(S + dS)$ and new knowledge portion dS. It is assumed that the components dI and dS are closely related to the elementary meanings, introduced in [1].

In [1] you also can find charts of the ontology graphs of scientific publications design and their processing. On the basis of (X, R) of the ontology graph developed a suitable ontology graph of the scientific publication. Publications formal XML/RDF description developed with the use of the Protégé 5.5.0 instruments.

Elementary meanings created out of the sentences of the text of the article, depending on the complexity of the sentence and its contents (in accordance with the syntax of the Ukrainian language) are formed like this:

- Complex sentences split into simple sentences.
- Simple and complicated sentences split into simple sentences that have subject, predicate and direct object.
- Simple two-member sentence represented as subject (S), predicate (P) and object (O).
- The elementary meaning is a certain equivalent of the RDF triple of the Semantic Web.

On the Figure 2 and Figure 3, you can see mentioned (X, R) ontology graph, and ontology graph of scientific publication, created with the Protégé 5.5.0 instruments. This ontology graphs describe "article 1", they are used to show examples of SPARQL queries and results too. English version of "article 1" was printed in Journal of Automation and Information Sciences, vol 50, 2018, Issue 10, PP. 1-17.

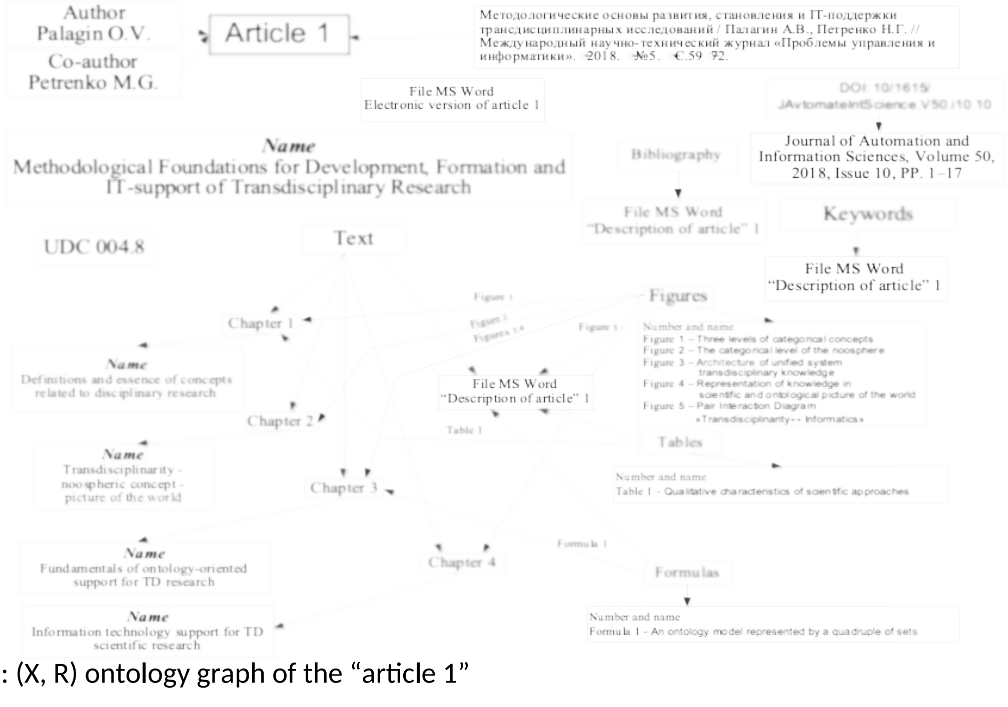
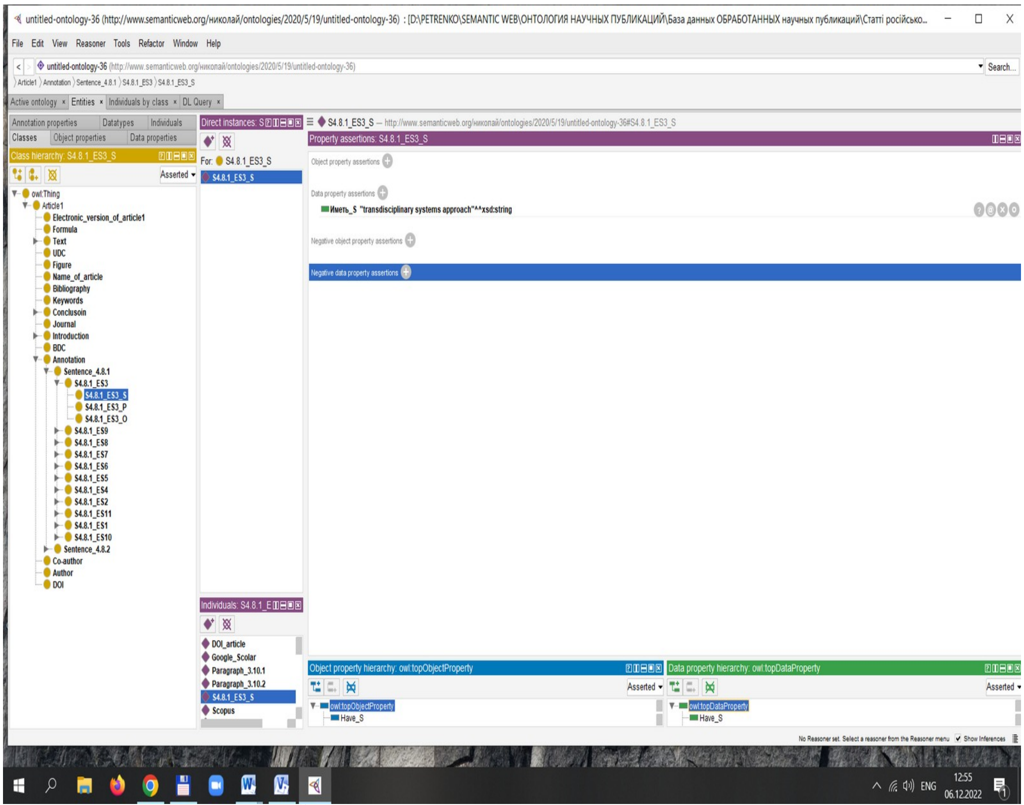

Author
Palagin O.V.
Co-author
Petrenko M.G.
Article 1
Методологические основы развития, становления и IT-поддержки трансдисциплинарных исследований / Палагин А.В., Петренко Н.Г. // Международный научно-технический журнал «Проблемы управления и информатики». 2018. №5. С.59 72.
File MS Word
Electronic version of article 1
DOI: 10/1615/JAvtomateInfScience.V50.i10.10
Journal of Automation and Information Sciences, Volume 50, 2018, Issue 10, PP. 1–17
Name
Methodological Foundations for Development, Formation and IT-support of Transdisciplinary Research
Bibliography
File MS Word
"Description of article" 1
Keywords
File MS Word
"Description of article" 1
UDC 004.8
Text
Figures
Figure 1
Figure 2
Figures 3,4
Figure 5
Chapter 1
Number and name
Figure 1 – Three levels of categorical concepts
Figure 2 – The categorical level of the noosphere
Figure 3 – Architecture of unified system transdisciplinary knowledge
Figure 4 – Representation of knowledge in scientific and ontological picture of the world
Figure 5 – Pair Interaction Diagram «Transdisciplinarity↔ Informatics»
Name
Definitions and essence of concepts related to disciplinary research
File MS Word
"Description of article" 1
Chapter 2
Table 1
Tables
Name
Transdisciplinarity - noospheric concept - picture of the world
Chapter 3
Number and name
Table 1 - Qualitative characteristics of scientific approaches
Name
Fundamentals of ontology-oriented support for TD research
Chapter 4
Formula 1
Formulas
Name
Information technology support for TD scientific research
Number and name
Formula 1 - An ontology model represented by a quadruple of sets

Figure 2: (X, R) ontology graph of the "article 1"



untitled-ontology-36 (http://www.semanticweb.org/николай/ontologies/2020/5/19/untitled-ontology-36) : [D:\PETRENKO\SEMANTIC WEB\ОНТОЛОГИЯ НАУЧНЫХ ПУБЛИКАЦИЙ\База данных ОБРАБОТАННЫХ научных публикаций\Статті російсько...
File Edit View Reasoner Tools Refactor Window Help
untitled-ontology-36 (http://www.semanticweb.org/николай/ontologies/2020/5/19/untitled-ontology-36)
Article1 〉 Annotation 〉 Sentence_4.8.1 〉 S4.8.1_ES3 〉 S4.8.1_ES3_S
Active ontology × Entities × Individuals by class × DL Query ×
Annotation properties Datatypes Individuals
Classes Object properties Data properties
Class hierarchy: S4.8.1_ES3_S
Asserted
owl:Thing
Article1
Electronic_version_of_article1
Formula
Text
UDC
Figure
Name_of_article
Bibliography
Keywords
Conclusoin
Journal
Introduction
BDC
Annotation
Sentence_4.8.1
S4.8.1_ES3
S4.8.1_ES3_S
S4.8.1_ES3_P
S4.8.1_ES3_O
S4.8.1_ES9
S4.8.1_ES8
S4.8.1_ES7
S4.8.1_ES6
S4.8.1_ES5
S4.8.1_ES4
S4.8.1_ES2
S4.8.1_ES11
S4.8.1_ES1
S4.8.1_ES10
Sentence_4.8.2
Co-author
Author
DOI
Direct instances: S
For: S4.8.1_ES3_S
S4.8.1_ES3_S
Individuals: S4.8.1_E
DOI_article
Google_Scolar
Paragraph_3.10.1
Paragraph_3.10.2
S4.8.1_ES3_S
Scopus
S4.8.1_ES3_S — http://www.semanticweb.org/николай/ontologies/2020/5/19/untitled-ontology-36#S4.8.1_ES3_S
Property assertions: S4.8.1_ES3_S
Object property assertions
Data property assertions
Negative object property assertions
Negative data property assertions
Object property hierarchy: owl:topObjectProperty
owl:topObjectProperty
Have_S
Data property hierarchy: owl:topDataProperty
owl:topDataProperty
Have_S
No Reasoner set. Select a reasoner from the Reasoner menu
Show Inferences
ENG
12:55
06.12.2022

Иметь_S "transdisciplinary systems approach"^^xsd:string

Figure 3: ontology graph of “article 1” in Protégé 5.5.0

The main stage of user tasks is broken down into three variants of the OrC architecture - A, B and C. These variants have different functional capacities. A – Least powerful (organized as a local endpoint on the user PC). B – Medium power (organized as a remote endpoint based on the Apache Jena Fuseki server). C – most powerful (organized as a remote endpoint, which is implemented using the original software). We can see that different variants of the OrC are suitable for different purposes. A – For a single user in a local network with a knowledge engineer (KE), in this scenario the user can form queries and get answers, working with only one science publication at a time. B – for multiple users of the same research unit. C – for users of the whole institute. For option B it is already possible to form a single query to retrieve structured information from several articles simultaneously, which cannot be done with popular search engines.

This material will focus on describing processes using UML-diagrams for variant B, phase 1 (B1).

2. Architectural and structural organization of the OrC DBSP (variant B, phase 1)

In this variant, the OrC functions as a remote endpoint based on the Apache Jena Fuseki and consists of three phases: phase 1 – processing of user SPARQL queries; phase 2 – multimedia visualization of user query results, or creation and use of conceptual and figurative structures of the subject area; phase 3 – manipulation of elementary meanings with purpose of gaining new knowledge.

Figure 4 shows a generalized diagram for the OrC B1 variant.

First, the knowledge engineer downloads the appropriate files and deploys Apache Jena Fuseki as a remote endpoint [16, 17]. He then uploads scientific publications to the server in the form of RDF graphs; this data is generated during the preparation stage.

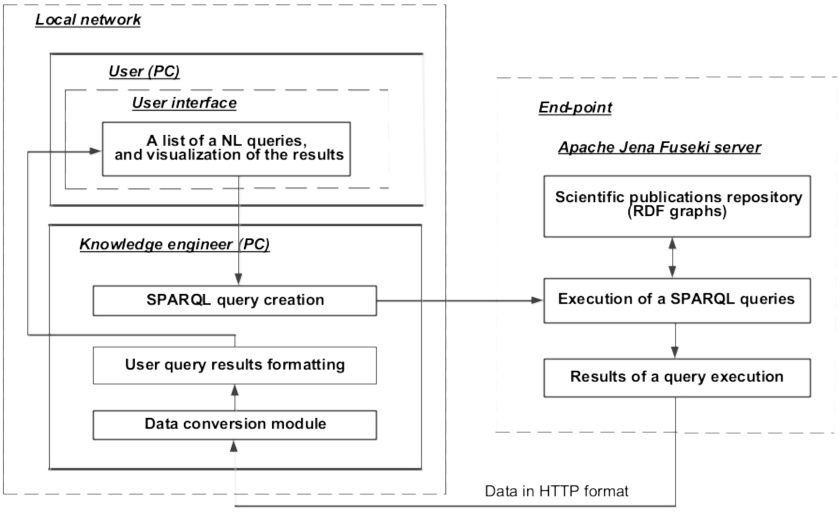

Local network
User (PC)
User interface
A list of a NL queries, and visualization of the results
Knowledge engineer (PC)
SPARQL query creation
User query results formatting
Data conversion module
End-point
Apache Jena Fuseki server
Scientific publications repository (RDF graphs)
Execution of a SPARQL queries
Results of a query execution
Data in HTTP format

Figure 4: Generalized diagram of the OrC DBSP

The user sees in his interface a list of possible natural language queries. He can select any query from this list one by one, the selected query is sent over the network to the knowledge engineer module. Systematically, the user specifies the information with which he works. It is possible to select a subset of the articles that are used for the search, this feature is useful if it is not necessary to search the entire database.

The following are examples of queries in natural language (NL).

3. Basic user queries

The researcher's database contains N scientific articles published in popular scientific journals. The serial numbers of scientific publications N can be described as follows:

$$\mathbf{N} = 1, 2, ..., m_1, ..., m_2, ..., m_k, ...$$

The serial numbers of scientific publications (in this case we are dealing with articles) serve as arguments for queries. The data are organized so that the author of a scientific publication is the first co-author in the publication; otherwise, the owner of the database is the author.

1. Show the titles of the articles on the subject of "transdisciplinarity".
2. Show the titles of the articles on the subject of "ontological".
3. Show the abstracts of the articles $m_1, \ldots, m_2, \ldots, m_k, \ldots$
4. Show the keywords of the articles $m_1, \ldots, m_2, \ldots, m_k, \ldots$
5. Show the titles of all **N** articles:
6. Show the titles of all **N** articles in the order of the date of publication;
7. Show the titles of all **N** articles without co-authors.
8. Show the titles of the articles $m_1, \ldots, m_2, \ldots, m_k, \ldots$, where m_1, m_2, m_k are user-defined query arguments.
9. Show the full names of the co-authors of the articles $m_1, \ldots, m_2, \ldots, m_k, \ldots$
10. Show the names of the chapters of the articles $m_1, \ldots, m_2, \ldots, m_k, \ldots$

…

4. UML-diagrams of the OrC functioning for variant B1

Now let us discuss UML-diagrams that reveal the core OrC functions for variant B1.

Figure 5 is the use case diagram, Figure 6 is the class diagram, Figure 7 is the components diagram, and Figure 8 is the sequence diagram.

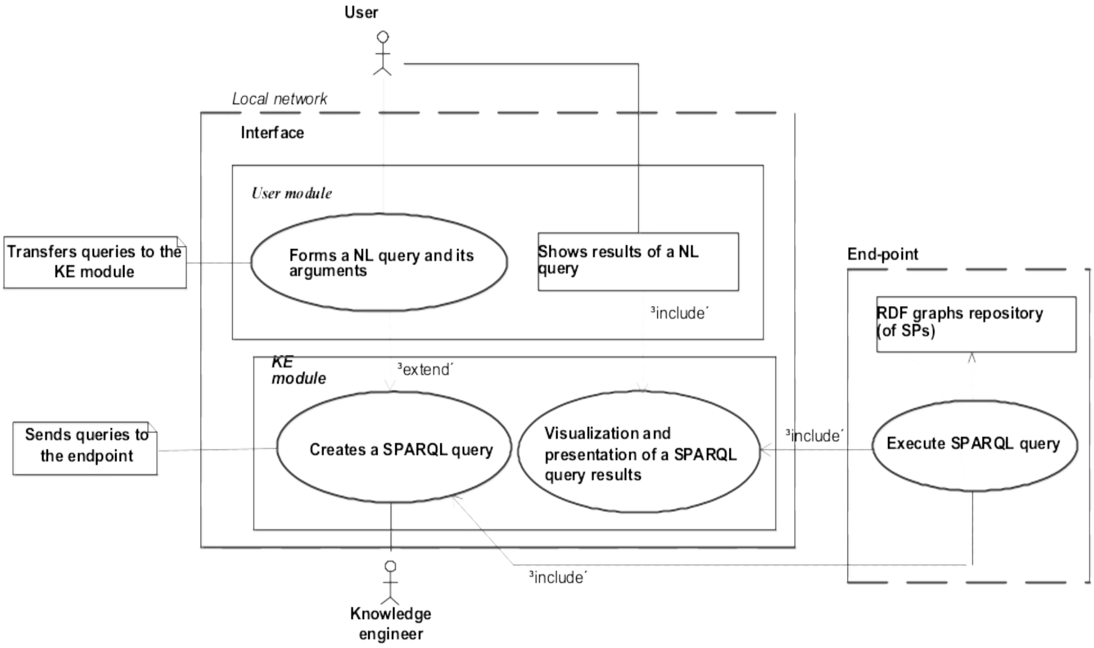

User
Local network
Interface
User module
Transfers queries to the KE module
Forms a NL query and its arguments
Shows results of a NL query
³include´
End-point
RDF graphs repository (of SPs)
KE module
³extend´
Sends queries to the endpoint
Creates a SPARQL query
Visualization and presentation of a SPARQL query results
³include´
Execute SPARQL query
³include´
Knowledge engineer

Figure 5: Use case diagram of the OrC DBSP

A certain number of researchers are connected to a local area network (LAN), which is managed by a knowledge engineer. We will look at the operation of the system for one user, for other users the process is organized in a similar way.

A general interface module functions on the researcher's personal computer (PC). The interface displays all queries in natural language. The researcher can select a single query with the required arguments. Another interface element shows the results of the query.

The other part of the system contains a knowledge engineer module. This module generates a SPARQL-query from a NL-query, and sends it over HTTP protocol to the end-point. The Apache Jena Fuseki server executes the SPARQL-query and sends the result back to the knowledge engineer module.

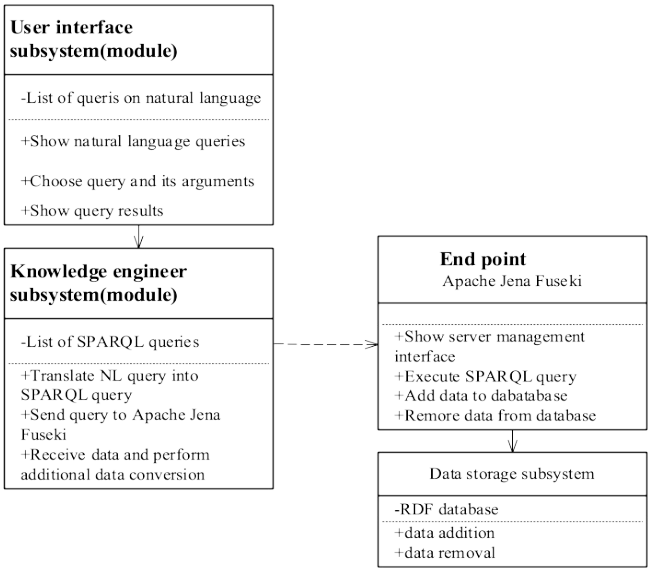

User interface subsystem(module)
-List of queris on natural language
+Show natural language queries
+Choose query and its arguments
+Show query results
Knowledge engineer subsystem(module)
-List of SPARQL queries
+Translate NL query into SPARQL query
+Send query to Apache Jena Fuseki
+Receive data and perform additional data conversion
End point
Apache Jena Fuseki
+Show server management interface
+Execute SPARQL query
+Add data to dabatabase
+Remore data from database
Data storage subsystem
-RDF database
+data addition
+data removal

Figure 6: Class diagram of the OrC DBSP

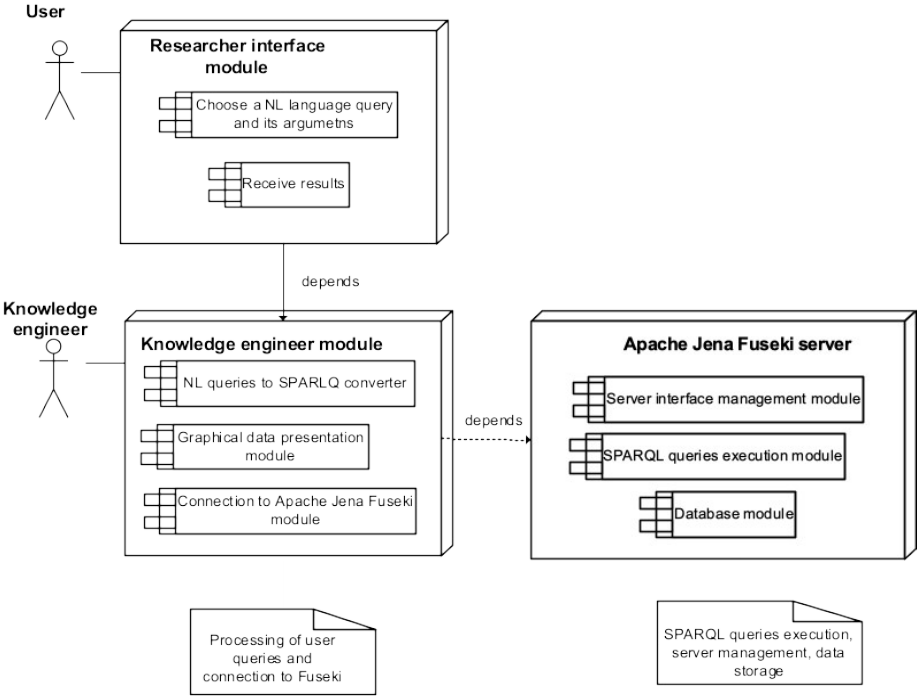

User
Researcher interface module
Choose a NL language query and its argumetns
Receive results
depends
Knowledge engineer
Knowledge engineer module
NL queries to SPARLQ converter
Graphical data presentation module
Connection to Apache Jena Fuseki module
depends
Apache Jena Fuseki server
Server interface management module
SPARQL queries execution module
Database module
Processing of user queries and connection to Fuseki
SPARQL queries execution, server management, data storage

Figure 7: Components diagram of the OrC DBSP

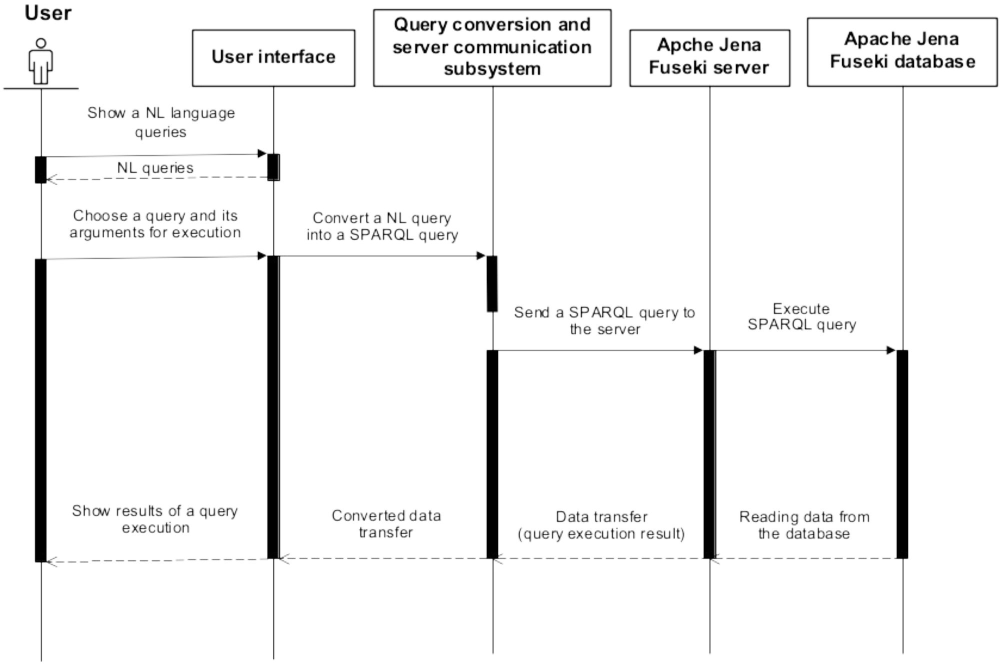

User
User interface
Query conversion and server communication subsystem
Apche Jena Fuseki server
Apache Jena Fuseki database
Show a NL language queries
NL queries
Choose a query and its arguments for execution
Convert a NL query into a SPARQL query
Send a SPARQL query to the server
Execute SPARQL query
Show results of a query execution
Converted data transfer
Data transfer (query execution result)
Reading data from the database

Figure 8: Sequence diagram OrC DBSP

The sequence of the generation and processing of the user queries is shown in detail in Fig. 5 to Fig. 8.

5. Examples of SPARQL-queries and their results

It is important to note that the diagrams do not show the process of selecting arguments and converting them into article numbers in the database.

NL-query.

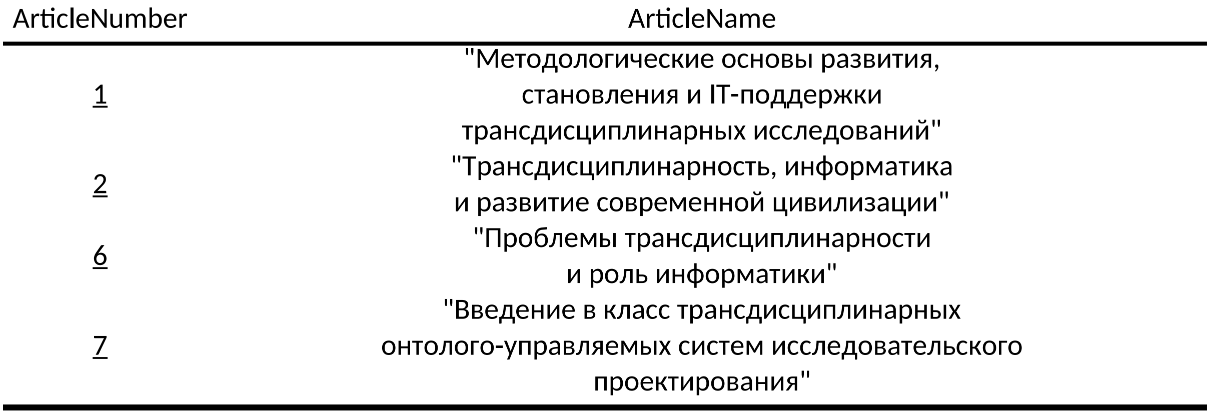
Show the titles of the articles on the subject of “transdisciplinarity”.

SPARQL-query.

```
PREFIX : <http://www.semanticweb.org/николай/ontologies/2020/5/19/untitled-ontology-36#>

SELECT ?ArticleNumber ?ArticleName
{
  GRAPH ?номер_статті {?s1 :Название_статьи ?ArticleName.
    FILTER REGEX(?ArticleName, "трансдисципл", "i")}
    {bind(strafter(str(?номер_статті),str(:)) as ?ArticleNumber)}
}
```

Query results.



ArticleNumber	ArticleName
1	"Методологические основы развития, становления и IT-поддержки трансдисциплинарных исследований"
2	"Трансдисциплинарность, информатика и развитие современной цивилизации"
6	"Проблемы трансдисциплинарности и роль информатики"
7	"Введение в класс трансдисциплинарных онтолого-управляемых систем исследовательского проектирования"

NL-query.

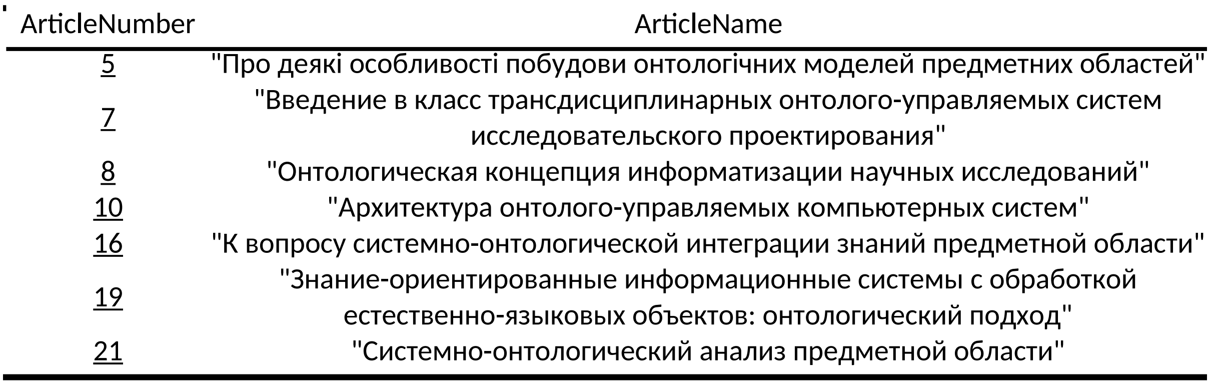
Show the titles of the articles on the topic of “ontological”.

SPARQL-query.

```
PREFIX : <http://www.semanticweb.org/николай/ontologies/2020/5/19/untitled-ontology-36#>

SELECT DISTINCT ?ArticleNumber ?ArticleName
{
  GRAPH ?номер_статті {?s1 :Название_статьи ?ArticleName.
    FILTER REGEX(?ArticleName, "онтолог", "i")}
    {bind(strafter(str(?номер_статті),str(:)) as ?ArticleNumber)}
}
```

Query results.



ArticleNumber	ArticleName
5	"Про деякі особливості побудови онтологічних моделей предметних областей"
7	"Введение в класс трансдисциплинарных онтолого-управляемых систем исследовательского проектирования"
8	"Онтологическая концепция информатизации научных исследований"
10	"Архитектура онтолого-управляемых компьютерных систем"
16	"К вопросу системно-онтологической интеграции знаний предметной области"
19	"Знание-ориентированные информационные системы с обработкой естественно-языковых объектов: онтологический подход"
21	"Системно-онтологический анализ предметной области"

NL-query.

Show the annotations of the articles 1, 2, 7.

SPARQL-query.

```
PREFIX : <http://www.semanticweb.org/николай/ontologies/2020/5/19/untitled-ontology-36#>

SELECT ?ArticleNumber ?ArticleName (group_concat(?анотація) as ?Abstract)
FROM NAMED <http://test.ulif.org.ua:51089/articles/data/article1>
FROM NAMED <http://test.ulif.org.ua:51089/articles/data/article2>
FROM NAMED <http://test.ulif.org.ua:51089/articles/data/article7>
{
  GRAPH ?номер_статті {?s1 :Название_статьи ?ArticleName.
                {:Аннотация :Иметь_Предложение  ?речення}
       {?речення :Иметь_Текст ?анотація}
       {bind(strafter(str(?номер_статті),str(:)) as ?ArticleNumber)}
       }
}
group by ?ArticleNumber ?ArticleName
```

Query results.

ArticleNumber	ArticleName	Abstract
1	"Методологические основы развития, становления и IT-поддержки трансдисциплинарных исследований"	"Разработаны основы методологии трансдисциплинарного системного подхода к постановке и выполнению научных исследований и сложных прикладных проектов с акцентом на их IT-поддержку с использованием методов и средств искусственного интеллекта, в частности онтологического инжиниринга. ...
2	"Трансдисциплинарность, информатика и развитие современной цивилизации"	"Перспективы и проблемы развития человеческой цивилизации всегда волновали общество. ...
7	"Введение в класс трансдисциплинарных онтолого-управляемых систем исследовательского проектирования"	"Рассмотрен класс систем исследовательского проектирования, основанных на использовании парадигм трансдисциплинарности, онтологического управления и целенаправленного развития. ...

NL-query.

Show the keywords of the articles 1, 2, 7.

SPARQL-query.

```
PREFIX : <http://www.semanticweb.org/николай/ontologies/2020/5/19/untitled-ontology-36#>

SELECT ?ArticleNumber ?ArticleName ?KeyWords
FROM NAMED <http://test.ulif.org.ua:51089/articles/data/article1>
FROM NAMED <http://test.ulif.org.ua:51089/articles/data/article2>
FROM NAMED <http://test.ulif.org.ua:51089/articles/data/article7>
{
      GRAPH ?номер_статті { ?s1  :Название_КС ?KeyWords OPTIONAL
   {?s2 :Название_статьи ?ArticleName}}
   {bind(strafter(str(?номер_статті),str(:)) as ?ArticleNumber)}
}
```

Query results.

ArticleNumber	ArticleName	KeyWords

7	"Введение в класс трансдисциплинарных онтолого-управляемых систем исследовательского проектирования"	"трансдисциплинарность, онтологическое управление, виртуальные структуры (парадигма), развивающиеся системы, ноосферогенез, ноосфера, научная картина мира, трансдисциплинарный подход (знания), кластеры конвергенции, онтологический подход, онтологическая концепция, формальная онтология, формула Брукса, интеллектуальные ИС, трансдисциплинарные онтолого-управляемые ИС, исследовательское проектирование, персональные базы знаний, предметная область, GRID-сети"
2	"Трансдисциплинарность, информатика и развитие современной цивилизации"	"трансдисциплинарность, информатика, мониторинг, кластер конвергенции, компьютерная онтология, knowledge engineering, Единая национальная сеть информатизации, глобальная сеть трансдисциплинарных знаний.
1	"Методологические основы развития, становления и IT-поддержки трансдисциплинарных исследований"	"научная картина мира, информационная технология, развивающаяся информационная система, трансдисциплинарность, трансдисциплинарные исследования, трансдисциплинарные знания, кластер конвергенции, онтология, онтологическая концепция, онтолого-ориентиро-ванная поддержка."

NL-query.

Show the names of the chapters of the article 1.

SPARQL-query.

```
PREFIX : <http://www.semanticweb.org/николай/ontologies/2020/5/19/untitled-ontology-36#>

SELECT ?ArticleNumber ?ArticleName ?ArticleChapter ?ArticleChapterName
FROM NAMED <http://127.0.0.1:3030/articles/data/article1>
{
  GRAPH ?номер_статті {?s1 :Название_статьи ?ArticleName
                 OPTIONAL {?ArticleChapter :Название_раздела  ?ArticleChapterName} }
                                {bind(strafter(str(?номер_статті),str(:)) as ?ArticleNumber)}
}
```

Query results.

ArticleNumber	ArticleName	ArticleChapte	ArticleChapterName

		r	
1	"Методологические основы развития, становления и IT-поддержки трансдисциплинарных исследований"	:Раздел_1	"Определения и сущность понятий, связанных с дисциплинарными исследованиями"
1	"Методологические основы развития, становления и IT-поддержки трансдисциплинарных исследований"	:Раздел_2	"Трансдисциплинарность – ноосферная концепция – картина мира"
1	"Методологические основы развития, становления и IT-поддержки трансдисциплинарных исследований"	:Раздел_3	"Основы онтолого-ориентированной поддержки ТД-исследований"
1	"Методологические основы развития, становления и IT-поддержки трансдисциплинарных исследований"	:Раздел_4	"Информационные технологии поддержки ТД научных исследований"

6. Conclusion

The aim of our research was to develop an ontology-related complex for semantic processing of scientific data, which will allow the researcher to significantly increase the rate of extraction of necessary information (in the form of cognitive structures) from their own sources.

This paper presented and described the architectural and structural organization of the OrC, which includes a local area network of a user and knowledge engineer PCs, and a remote endpoint based on the Apache Jena Fuseki server. UML diagrams show the functioning of the OrC. Also shown are examples of user queries.

7. Further research

This study is still far from its final goal. As we have already explained, phases 2 and 3 need to be implemented. Algorithms for creating conceptual and figurative structures, algorithms for comparing and analyzing these structures with the further intention of constructing KD data, and algorithms for discovering new knowledge, including the algorithms according to the Brooks formula, need to be developed.

In future research, our team will develop original tools and facilities in order to optimize user queries and optimize the usability of the ontology-related complex.

References

[1] A. Palagin, N. Petrenko, Knowledge-oriented tool complex processing databases of scientific publications. Control systems and computers. 2020. №5. p. 17–33. doi.org/10.15407/csc.2020.05 [Accessed: 22 June 2022].

[2] A. Palagin, N. Petrenko, V. Velychko, K. Malakhov, Development of formal models, algorithms, procedures, engineering and functioning of the software system “Instrumental complex for ontological engineering purpose”. In: Proceedings of the 9th International Conference of Programming UkrPROG. CEUR Workshop Proceedings 1843. Kyiv, Ukraine, May 20-22, 2014. [Online] Available from: http://ceur-ws.org/Vol-1843/221-232.pdf [Accessed: 23 June 2022].

[3] A. Palagin, S. Kryvyy, N. Petrenko, Ontological methods and means of processing subject knowledge. Lugansk: V.I. Dal East Ukr. Nac. University. [online] Available at: <http://www.aduis.com.ua/Monography.pdf> [Accessed: 20 June 2022].

[4] Designing and program implementation of the subsystem for creation and use of the ontological knowledge base of the scientific employee publications. Problems in programming. 2017. №2. p. 72–81. doi.org/10.15407/pp2017.02.072.

[5] A. Palagin, V. Velychko, K. Malakhov, O. Shchurov, Distributional semantic modeling: a revised technique to train term/word vector space models applying the ontology-related approach. In: Proceedings of the 12th International Scientific and Practical Conference of Programming UkrPROG 2020. CEUR Workshop Proceedings 2866. Kyiv, Ukraine, September 15-16, 2020. [Online] Available from: http://ceur-ws.org/Vol-2866/ceur_342-352palagin34.pdf [Accessed: 20 June 2022].

[6] A. Palagin, N. Petrenko, (2018) Methodological Foundations for Development, Formation and IT-support of Transdisciplinary Research. Journal of Automation and Information Sciences. 50(10). p. 1-17. doi.org/10.1615/JAutomatInfScien.v50.i10.10.

[7] S. Kotlyk, (Ed.) New Information Technologies, Simulation and Automation. Iowa State University Digital Press. doi.org/10.31274/isudp.2022.121.

[8] A. Palagin, Transdisciplinarity problems and the role of informatics. Cybernetics and Systems Analysis/ International Theoretical Science Journal. 2013, № 5 – p.3–13. doi.org/10.1007/s10559-013-9551-y.

[9] A. Palagin, Architecture of ontology-driven computer systems. Cybernetics and Systems Analysis/ International Theoretical Science Journal. 2006, № 2 – p. 254–264. doi.org/10.1007/s10559-006-0061-z.

[10] A. Palagin, An Introduction to the Class of the Transdisciplinary Ontology-controled Research Design Systems. Control systems and computers. 2016. № 6. p. 3–11. doi.org/10.15407/usim.2016.06.003.

[11] A. Palagin, A. Kurgaev, Interdisciplinary scientific research: optimization of system and information support. (2009). Bulletin of the National Academy of Sciences of Ukraine. 2009, № 3, p.14–25. [Online] Available from: ftp://ftp.nas.gov.ua/akademperiodyka/Downloads/ Visnyk_NANU/downloads/2009/3/st3.pdf [Accessed: 23 June 2022].

[12] A. Palagin, N. Petrenko, S. Kryvyy, On the construction of knowledge-oriented computer systems for scientific research. Control systems and computers. 2015. №2. p. 64–73. [Online] Available from: http://usim.org.ua/arch/2015/2/7.pdf [Accessed 22 June 2022].

[13] G. Booch, J. Rumbaugh, I. Jacobson, The Unified Modeling Language User Guide. Reading, MA, 2005. 475 p.

[14] D. Schmuller, Teach Yourself UML in 24 Hours, Complete Starter Kit. M.: Williams, 2005. 416 p.

[15] A. Leonenkov, Tutorial UML 2. St. Petersburg: BHV-Petersburg, 2007. 576 p. ISBN 978-5-94157-878-8.

[16] https://jena.apache.org/documentation/fuseki2/(date of access: 23 June 2022).

[17] B. Ducharme, Learning SPARQL. Querying and Updating with SPARQL 1.1 (Second edition). O’Reilly Media, August 2013. 367p.