

*Distributed information systems that support scientific and educational activities can work with various information systems. The main goal of creating a distributed information system supporting scientific and educational activities is to accelerate the pace and improve the quality of information exchange in the scientific environment.*

*The paper considers technological methods for constructing models of information systems designed to support scientific and educational activities. The model under consideration is that the developed model of an information system for working with scientific materials should solve the problems of long-term storage of information, organizing data search by attributes, accumulating and replacing metadata.*

*Based on the analysis of typical scenarios of information servers, the tasks that should be solved when organizing an access control system for distributed information resources are formulated. Within the framework of this technology, three access control models are discussed, which differ in the degree of integration of information server functions with the Z39.50 technologies.*

*The creation and support of distributed information systems and electronic libraries that integrate heterogeneous information resources and operate in various software and hardware environments require special approaches to managing these systems. If the resources or data themselves can be managed locally, even for distributed information systems, then the task of managing access to distributed resources cannot be solved within the framework of local administration. The justification of the last thesis can be seen when considering typical scenarios of the information server, which we will describe below*

*Keywords: access, information systems, distributed information systems, storage, search, technology, model, Z39.50 protocol*

UDC 004.7:004.75

DOI: 10.15587/1729-4061.2021.239540

# DEVELOPMENT OF QUERIES USING THE Z39.50 PROTOCOL IN DISTRIBUTED INFORMATION SYSTEMS TO SUPPORT SCIENTIFIC AND EDUCATIONAL ACTIVITIES

**Sandugash Serikbayeva**

*Corresponding author*

Doctoral Student, Teacher\*

E-mail: inf\_8585@mail.ru

**Jamalbek Tussupov**

Doctor of Physical and Mathematical Sciences, Professor,

Head of Department\*

**Madina Sambetbayeva**

PhD, Associate Professor, Leading Researcher\*\*

**Gaukhar Muratova**

Associate Professor, Dean of Faculty

Faculty of Information Technologies, Automation and

Telecommunications\*\*\*

**Mukhtar Makhonov**

Candidate of Technical Sciences, Associate Professor

Department of Engineering

Eurasian Technological University

Tole bi str., 109A, Almaty, Republic of Kazakhstan, 050000

**Gaukhar Borankulova**

PhD, Associate Professor, Head of Department

Department of Information Systems\*\*\*

**Aigerim Yerimbetova**

PhD, Associate Professor, Leading Researcher\*\*

Department of Software Engineering

Institute of Automation and Information Technologies

Satbayev University

Satbayev str., 22a, Almaty, Republic of Kazakhstan, 050013

\*Department of Information Systems

L. N. Gumilyov Eurasian National University

Satpayev str., 2, Nur-Sultan, Republic of Kazakhstan, 010000

\*\*Institute of Information and Computational Technologies

Shevchenko str., 28, Almaty, Republic of Kazakhstan, 050010

\*\*\*M. Kh. Dulaty Taraz Regional University

Tole bi str., 40, Taraz, Republic of Kazakhstan, 080000

Received date 23.08.2021

**How to Cite:** Serikbayeva, S., Tussupov, J., Sambetbayeva, M., Muratova, G., Makhonov, M., Borankulova, G., Yerimbetova, A. (2021).

Accepted date 18.10.2021

Development of queries using the Z39.50 protocol in distributed information systems to support scientific and educational activities.

Published date 29.10.2021

Eastern-European Journal of Enterprise Technologies, 5 (2 (113)), 66–79. doi: <https://doi.org/10.15587/1729-4061.2021.239540>

## 1. Introduction

The problem of access to information (including computing) resources is one of the main problems that arise in the activities of the scientific and educational community. Currently, there is a transition to a distributed scheme for creat-

ing and maintaining scientific and information resources, on the one hand, and the desire for virtual unity by providing free access to any resources in the network through a limited number of "access points", on the other.

It is obvious that with the growth of the number of resources, various problems associated with their support and

management invariably arise. First of all, these problems are caused by the need to duplicate data used by different resources. If the stored data is updated frequently enough, logical contradictions between the data on different resources periodically arise, which can become a source of errors and failures in the functioning of resources. In addition, the very fact of duplication greatly increases the work of system administrators associated with maintaining resources. Constantly acting factors in the formation of a single (virtual) information space of the organization are:

- hierarchy of information systems and resources;
- heterogeneity of resources and software and hardware environments combined in a single network operating space;
- distribution of information infrastructure elements.

The realization of the need to integrate heterogeneous information resources led to the creation of integrated (unified) scientific information systems (SIS), which would allow establishing links between heterogeneous documents, organizing unified catalogs of documents, as well as creating specialized search systems. The main problem associated with the functioning of integrated distributed information systems is widely known – it is practically a non-functioning information updating system. It is almost impossible to solve this problem by administrative methods. Note that the effective operation of information resources is possible only if they are constantly supported by the authors.

The only way to solve this problem is to integrate the data of local information and reference systems existing in the organization within an integrated distributed information system and give these systems the functions of global (corporate) authentication and authorization of users to access information resources.

Distributed information systems that support scientific and educational activities can work with various information systems. The main goal of creating a distributed information system supporting scientific and educational activities is to accelerate the pace and improve the quality of information exchange in the scientific environment. One of the most pressing issues is the division of the unified compatibility of the information system and the work on the systematization of information resources into professional areas. These can be scientific articles, scientific documents, electronic collections, ontological descriptions, data sets, logical descriptions, and so on. Semantic connections between information resources increase their value and provide additional opportunities for searching and identifying information.

---

## 2. Literature review and problem statement

---

The paper [1] discusses the issues of creating a technological model of an integrated distributed library information system that combines the digitized book collection and scientific works of the Kazakhstan Engineering and Technology University and some research institutes located in the Academic City of Almaty.

The work [2] describes a model of an integrated distributed library information system that combines the digitized book collection and scientific works of the Kazakhstan University of Engineering and Technology and some research institutes located in the Academic City of Almaty. The possibilities and needs of all participants of the scientific and educational cluster for building an optimal architecture of a distributed information system are analyzed. The descrip-

tion is given and the results of the implementation of subsystems of the digital repository of the integrated distributed library information system of the Academic City of Almaty are presented.

In [3], the development of modern society is characterized by an increasing volume and rapid obsolescence of scientific information. To increase the effectiveness of scientific research, scientists need access to information about the results of research carried out in the field of interest. Therefore, any scientific research usually begins with the search for scientific information about research in this field, but the search for necessary information in an ever-increasing volume of articles, books, monographs, reports, patents is becoming increasingly difficult. Scientists have to spend a lot of time searching and processing information that allows them to quickly get acquainted with the results of other studies and eliminate their duplication. The paper defines information systems designed to support scientific and educational activities from the point of view of scientific communication. The task, subject area, subjects, objects, the main functionality of the information system are defined, a list of the main types of information resources is provided. The paper analyzes the functional requirements for such systems.

In [4], the task of developing a technology for supporting large information repositories and organizing delimited user access to this information, providing a service for both managing these objects and organizing access to these objects, is solved. The solution of the problem will allow you to create a conceptual model with the allocation of basic entities among information objects and the establishment of links between them. It will also allow us to develop technical documentation reflecting the results of the first stage of creating an information system: solving problems of syntactic and technical interoperability, developing a single interface, interacting with users, etc. To implement search functions outside of graphical interfaces, support for special network services and query languages is required. Ideally, all IS should support a single search profile and a single query language.

The paper [5] solves the problem of developing a technology for supporting large information repositories and organizing delimited user access to this information, providing a service for both managing these objects and organizing access to these objects. Solving the problem will allow you to create a conceptual model with the allocation of basic entities among information objects and the establishment of links between them. It will also allow the development of technical documentation reflecting the results of the first stage of the creation of an information system: solving problems of syntactic and technical interoperability, developing a single interface, interacting with users, etc.

In existing EL developments, as a rule, search and access to information are provided only through visual graphical interfaces. The task of the subsystem for integrating various electronic resources is to provide other subsystems with a single interface for accessing information stored in the system's data sources. That is, any resource should be cataloged in a standard way, provided with metadata, access rules, and a unique identifier.

To implement search functions outside of graphical interfaces, support for special network services and query languages is required. Ideally, all IS should support a single search profile and a single query language.

The paper [6] describes the main tasks of creating a model of a distributed information system, supporting scientific and educational activities, the functionality of the model, the concept of metadata, and the requirements for the metadata profile. The task, subject area, subjects, objects, the main functionality of the information system are defined, a list of the main types of information resources is given. The paper analyzes the functional requirements for such systems.

The paper [7] describes the architecture of the system and the principles of integration with a digital repository, the rules for the representation and transformation of metadata. The main attention is paid to working with dictionaries of key terms that are used for systematization and classification of information resources and modeling links with facts.

The work [8] describes the issues related to the creation and support of distributed information and computing resources in the Siberian Branch of the Russian Academy of Sciences. The main technological approaches, the metadata structure and the use of GRID technologies are described.

In general, considering the scientific works of the above-mentioned authors, it is necessary to form the main points of view on the considered data and its systematization in the creation of a distributed information system that supports scientific and educational activities. Ways to create an abstract query for term dictionaries in the thesaurus using the Z39.50 protocol in a distributed information system that supports scientific and educational activities are considered.

The provision of access to information about heritage information systems is organized in various ways and using various technologies and protocols that often do not comply with existing standards.

Most of the existing information retrieval systems support completely different data storage structures, access methods and information presentation formats and, as a result, have their own user interface.

One of the important tasks in the development of distributed information and search systems for cultural heritage is the unification of access and integration of information resources. The goal is to combine the available information resources on cultural heritage into a single distributed information system with end-to-end search.

However, for heterogeneous and isolated information systems that do not support existing standards, there is a problem with the unification of data access and integration of information resources. All this makes it difficult to work with them in many ways, as well as further use the information provided by them.

The review of the above literature is the search for information on information abstraction data of distributed information systems that support scientific and educational activities. Based on the works of these authors, the Z39.50 protocol is used in scientific and educational activities, as a technology and protocol. The only technology currently that can solve the above problems is a technology based on international standards ANSI/NISO Z39.50.

---

### 3. The aim and objectives of the study

---

The aim of the study is that distributed information systems that support scientific and educational activities can work with various information systems. The purpose of the presented work is to organize abstract access to the thesaurus database using the Z39.50 protocol technology. Despite

the listed difficulties that arise when creating applications that access relational data through Z39.50, information systems built on this principle can combine the inherent versatility of Z39.50 in data exchange and the power of relational databases in storing and processing information. We hope that this material will be of interest both to developers of Z39.50 applications using relational databases and to a wide range of specialists in the field of open information systems.

To achieve this aim, the following objectives are set:

- exploring the capabilities of the Z39.50 protocol;
- collection of theoretical material on the topic of abstract access to thesaurus;
- analysis of existing standards and thesaurus presentation data schemes, as well as search attributes designed for thesaurus search;
- development of an application that allows you to perform an abstract thesaurus search, as well as form a tree of terms.

---

## 4. Materials and methods

---

### 4.1. System of definitions, agreements, and rules

The main tasks for building a model of a distributed information system supporting scientific and educational activities, ensuring the functioning of the model, the concept of metadata for this system, and the requirements for the metadata profile are described. A description of the twelve requirements of a distributed information system that supports scientific and educational activities is given, systematizing them. Based on the proposed requirements, the architecture of a distributed information system supporting scientific and educational activities was developed and the structure was described.

The model parameters should take into account the specifics of the subject area, which can be expressed through the properties of semantic information resources (thematic composition, geographical and chronological coverage, level of processing, etc.) and physical nature (identification and location address, access method, access restrictions, etc.).

To manage, search and provide access to distributed information resources, it is proposed to use a special category of metadata – system metadata containing descriptions of the semantic and physical properties of resources and their structural elements.

System metadata provides information interfaces between data sources and the integrated information space. Through this, the functions of resource integration, navigation in the information space, search, and access to the necessary resources are implemented.

The main goal of creating a distributed information system supporting scientific and educational activities is to accelerate the pace and improve the quality of information exchange in the scientific environment. One of the most pressing issues is the unified integration of the information system and the division of work on the systematization of information resources into professional areas. These are scientific articles, scientific documents, electronic collections, ontological descriptions, data sets, logical descriptions, and so on. Semantic connections between information resources increase their value and provide additional opportunities for searching and identifying information. The thesaurus supports the Z39.50 protocol profile. In this regard, there is a need to develop an application for abstract access to the thesaurus. The application must use a single interface to search

through thesaurus with different data structure (schema) that supports the Z39.50 protocol profile.

The Z39.50 protocol defines the order of interaction between the client and the server, the procedures for searching and extracting information from databases, and the formats for presenting this information.

The Z39.50 protocol does not define data storage formats in specific databases, methods of indexing them, and procedures for the functioning of various DBMSs. The Z39.50 protocol also does not define user-client interaction interfaces.

In the ideology of Z39.50, within the same schema, all databases are the same, despite their physical differences in the DBMS used, fields, and query syntax. In Z39.50, the client can't determine under which DBMS the data extracted by him is stored. This seeming limitation makes a lot of sense, the essence of which will be clarified in the following sections. However, it can be noted that the client does not need such information, because he always works with the same query system and receives data in the same formats.

The constructed model of distributed information resources assumes a system of definitions, agreements, and rules that define methods of systematization and general mechanisms for integration and access to heterogeneous and geographically distributed information data.

#### 4. 2. Distributed information systems to support scientific and educational activities

The development of distributed information resources of an organization leads to the need to create an infrastructure for their integration into a single information system that provides transparent access to distributed information.

The development of global information and computing networks today leads to a change in the fundamental paradigms of working with information resources. The transition to distributed resources, the creation of an infrastructure for their integration into a single information system that provides transparent access to distributed information are relevant.

Therefore, the most important task related to the technology of working with information is to study ways to integrate distributed data sources and create a scientific reserve in the field of distributed information systems and databases to develop a technology that supports the creation and operation of large-scale information infrastructures based on virtual integration. This technology will allow creating global infrastructures from dozens and hundreds of heterogeneous databases and solving strategic tasks in the field of automation of various forms of distributed activities. A narrower goal is to develop principles and software tools for virtual integration of distributed data sources based on international standards and recommendations for creating large-scale information infrastructures designed to virtualize data access to various DBMS using common rules and policies [6].

The purpose of creating the system: to develop a system for distributed information systems to support scientific and educational activities, to create a program to study the capabilities of the Apache Solr platform for processing distributed data, which uses big data technologies.

A set of the most general functional requirements for the IP support of scientific and educational activities was identified.

- 1) collection of information resources;
- 2) the relevance of documents;
- 3) the relevance, completeness, reliability of the origin of documents;

- 4) the use of intelligent services for processing user requests;
- 5) knowledge extraction;
- 6) support for non-centralized information system architectures;
- 7) structuring of the information space;
- 8) the use of information classification in information search;
- 9) adaptive presentation of information;
- 10) historicity of information;
- 11) archive;
- 12) support for distribution.

In the conditions of working in a distributed environment, the following requirements are imposed on the IR support for scientific and educational activities:

- support of accepted metadata standards for data export and import;
- support of information exchange protocols with other information systems;
- support for the ability to link to internal resources both in user interfaces and at the system level.

The task of information systems is to store information and provide it to users in a convenient form. As a rule, such systems can be organized based on various technological solutions aimed at implementing a particular distribution paradigm. The distribution paradigm can be considered from the point of view of the architecture of information systems. Note that most information systems today are built on the principle of a three-tier architecture with the conditional division of links on clients, application servers, and database servers. Based on this, we can distinguish three main groups of distributed systems that implement the principle of distribution at the appropriate level.

The creation and support of distributed information systems and electronic libraries that integrate heterogeneous information resources and operate in various software and hardware environments require special approaches to managing these systems [7]. All information is stored in a DBMS based on the freely distributed PostgreSQL software; user and administrative interfaces are implemented based on Apache Solr. In the search area, Apache Solr has become the de-facto platform for creating production applications. Although Solr is designed to scale using a distributed, partitioned architecture, the platform is mainly designed around providing low-latency search for users.

To implement a system for processing large amounts of data, the task is to create a mock-up application. The application demonstrates the processing of big data distributed on several machines using the example of counting identical words in text datasets. Usually, the dataset has to be distributed across several different machines. Since these machines work on the network, the system administrator has to take into account all the complexities of network programming. One of the problems of working with large amounts of data is the difficulty of transferring them between servers for subsequent processing. Also, due to a large number of nodes, frequent failures of individual nodes are possible, so the issue of reliability is also very important. This problem can be solved using the Apache Solr technology.

System tasks:

- 1) collection, storage, and selection of unique publications from the internet space to the system database;
- 2) distribution of publications by topic: clustering, classification, definition of thematic combinations, ranking and filtering (by social spheres, regions, industries, etc.);

- 3) determination of information occasions;
- 4) calculation of the degrees of informative features of publication, such as collective use of purchased electronic literature catalogs, databases, and bibliographic publications;
- 5) identification of information trends.

To meet these requirements, it is necessary to create an infrastructure (an information service or a center) for the presentation and exchange of metadata – structured information about information resources and access rules to them. Currently, many information centers engaged in the collection and dissemination of metadata are actively interested in organizing interaction to exchange their existing funds. As a rule, such integration of funds is based on the development of a standard for the format for the presentation of metadata, simultaneously with the unification of arrays of normative reference information.

As part of the tasks set, the architecture of the information system was developed (Fig. 1) to systematize the resources of the electronic library, a multi-level DL architecture is used, consisting of a data warehouse, a repository, a metadata server, an application server, a dictionary, reference books, as well as a software implementation of the developed architecture deployed on existing hardware and put into operation.

Based on the described information system and a database of publications on information technologies, the following have been created:

- a detailed dictionary (thesaurus) of concepts and key terms in computer science and tools for its modification;
- thesaurus on information security as part of the thesaurus on computer science.

The system continues to develop both in terms of expanding functionality and adding new information resources.

The repository is an autonomous search engine and includes:

- electronic catalog data;
- indexing profile of electronic catalog entries;
- indexes for searching and viewing the electronic catalog;
- the repository can contain many electronic catalogs.

This model allows us to create various configurations of the electronic library catalog based on one or more repositories. Once created, the configuration can change dynamically during operation. It should be noted here that only one of the available data indexing profiles is used for the repository. This profile cannot be changed as long as there is at least one electronic catalog in the repository [5].

### 4.3. Models of distributed information systems to support scientific and educational activities

The Z39.50 standard defines a client/server type service and protocol for information retrieval. It specifies procedures and formats for the client to search in databases provided by the server, extract records from databases and perform other functions related to information search. The Z39.50 protocol defines interaction only between client and server information retrieval applications, it does not define interaction between the client and the end user. More precisely, the Z39.50 protocol does not define data storage formats in specific databases, ways of indexing them, and procedures for the functioning of various DBMS. It also does not define user-client interaction interfaces. Without going into the details of the protocol, we can say that the Z39.50 standard defines such rules for computer interaction that allow you to unify access to various databases. Thus, a user using only one client application can search for information in remote distributed databases with a very different structure and formats of information presentation. Two main features distinguish the Z39.50 protocol from other protocols. Firstly, it is an abstract model of information representation. In the ideology of Z39.50, within the same data schema, all databases are exactly the same, despite their physical differences in the DBMS used, fields and query syntax. In other words, the protocol provides an abstract model for presenting information at each stage of client-server interaction. In Z39.50, the client always works with the same query system and receives data in the same formats. The second feature is that the Z39.50 protocol fully provides session interaction between the client and the server. This feature is embedded in the protocol itself and is implemented in all its applications, whether it is a server system or a client program.

The purpose of this standard is to facilitate interaction between clients and servers in application systems in which the client searches and retrieves records from server databases. Databases may have different implementations: different systems may have different ways of storing data and different ways of accessing them. Therefore, when describing databases in Z39.50, a general abstract database model is used, to which each system can match its implementation. This allows different systems to interact using standard and commonly understood terms for the tasks of searching and extracting information from databases. Information retrieval and retrieval models will be discussed later in the third and fourth sections, respectively.

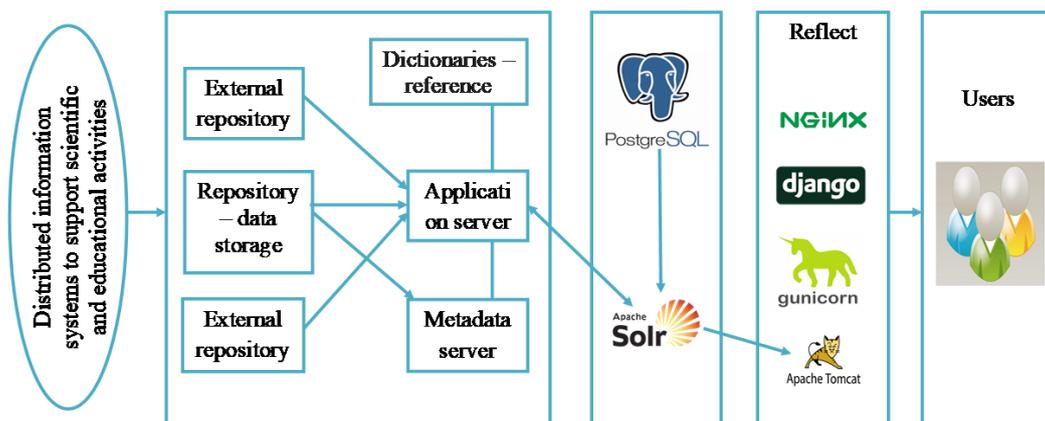


Fig. 1. The architecture of a distributed information system for supporting scientific and educational activities

In Z39.50, search queries are always formulated not to a real database, but some abstract one. This abstract database has no structure, it is characterized only by access points (search attributes). When a request is received from the client in the form of terms and search attributes, the server converts it into the syntax of a real database, and this procedure remains invisible to the client. With this approach to the search procedure, all databases become the same for the client if they support the same set of access points.

The database schema is a mutual agreement between the client and the server about the information contained in the database records, which allows you to further select some of this information following the specification of the element. The schema defines the abstract structure of the record. This is the primary element of the database schema, which is a tree of elements specified by tags from standard sets of tags (tagset). When applying an abstract record structure to a database record, an abstract database record is obtained. An abstract database record is an abstract representation of the information contained in a database record. To form an abstract record, you need to apply an abstract record structure defined in the schema to a database record.

To search for records, you need to specify a list of database servers and a list of database names to search for, as well as formulate a search query containing search criteria. In traditional systems using Z39.50, it is also necessary to know additional information to build a query: the syntax of the query language for each DBMS, as well as the structure, field names, and data types of each database. Fig. 2 shows the information retrieval model Z39.50.

It should be noted that Z39.50 is not a search engine. The Z39.50 client can send search results to one or more databases on remote systems simultaneously. The model allows clients to connect to each individual server by searching for the current contents of the database and getting the results directly from the source databases. But the web search engine is, in fact, a single information search engine that has an additional function of collecting resources from the Internet and performing a kind of indexing to make these resources searchable. In distributed and integrated access, searching for information from a single server is not too difficult, but it becomes more problematic if the search occurs from multiple databases on multiple servers. For Z39.50, it is also difficult to understand the requests and responses of the two systems. The lack of semantic interoperability has led to a loss of user confidence in the Z39.50 interface for information retrieval systems. These are a number of complex problems that Z39.50 researchers and developers have to face.

If the listed characteristics are different in a group of selected servers and databases, even a very simple query cannot be executed for the entire group. In Z39.50, this problem is solved by building a specific search model and standardizing its components.

In Z39.50, search queries are always formulated not to a real database, but some abstract one. This abstract database has no structure, it is characterized only by access points (search attributes). When a request is received from the client in the form of terms and search attributes, the server converts it into the syntax of a real database, and this procedure remains invisible to the client. With this approach to the search procedure, all databases become the same for the client if they support the same set of access points.

The sets of search attributes make up the class of objects Z39.50 {Z39.50 3} that are subject to standardization. At the moment, the sets of attributes presented in Table 1 are standardized.

Table 1

Attribute sets Z39.50

OID	Set	Comment
{Z39.50 3 1}	Bib-1	Bibliographic information
{Z39.50 3 2}	exp-1	Explain
{Z39.50 3 3}	ext-1	Advanced Service
{Z39.50 3 4}	ccl-1	Type-2 (ISO 8777) and Type-100 (Z39.58) requests
{Z39.50 3 5}et	gils	To search for GILS
{Z39.50 3 6}	stas	Scientific and technical information
{Z39.50 3 7}	collections-1	Navigation through electronic collections
{Z39.50 3 8}	cimi-1	Information on museum collections
{Z39.50 3 9}	geo-1	Spatial metadata
{Z39.50 3 10}	ZBIG	Biological information
{Z39.50 3 11}	util	Utilities
{Z39.50 3 12}	xd-1	Cross-domain set
{Z39.50 3 13}	Zthes	Navigating through thesaurus

To search for bibliographic information, a set of bib-1 attributes is used, as shown in Fig. 3. Some attribute sets (gils, geo-1) include the bib-1 set, so this set is the main one.

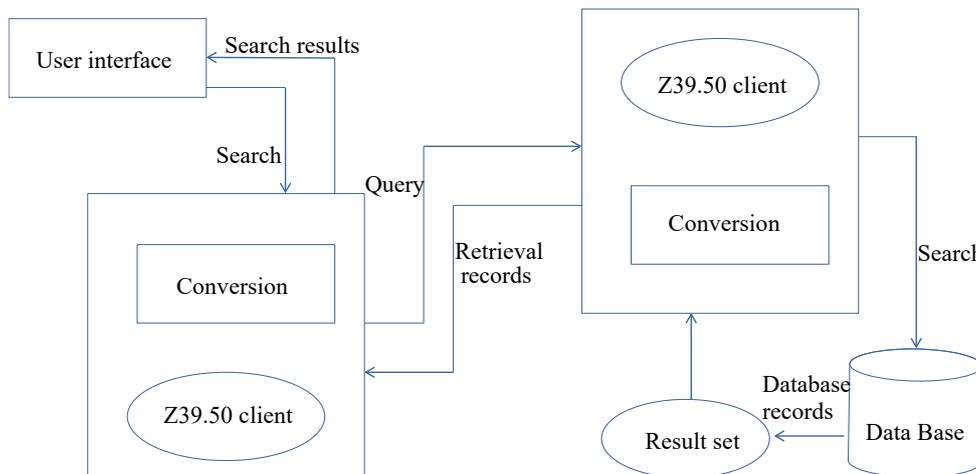


Fig. 2. Information retrieval model Z39.50

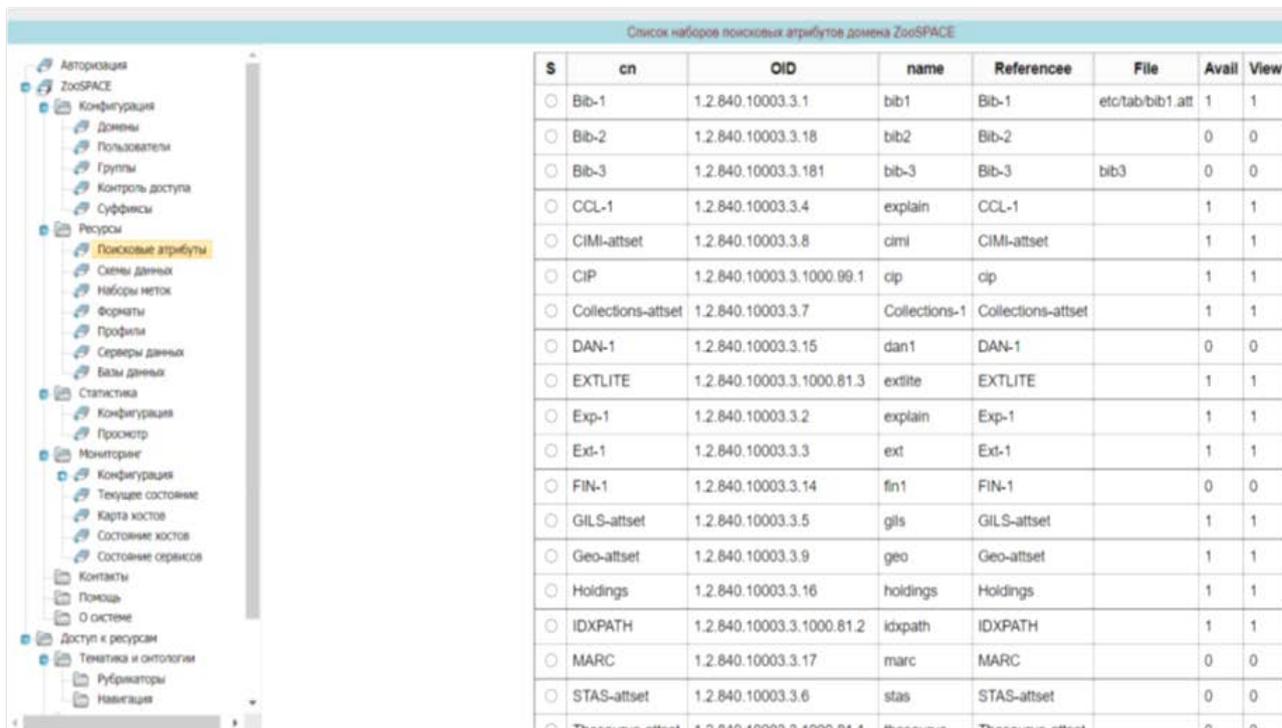


Fig. 3. Search attributes

The set of Bib-1 attributes [7] used for searching bibliographic information is the main one, its subsets are included in many other sets.

*Types of search attributes.*

The Bib-1 set includes six types of attributes with numbers 1–6: Use, Relation, Position, Structure, Truncation, and Completeness. When building a query, specifying search attributes in combination with a search term determines the search criteria. In each group, each attribute is defined by a unique numeric value, so you need to specify two numbers to specify the search attribute: type+value.

Attributes 1: Use (use attributes).

Attributes of this type indicate which semantic information the search term is associated with, i.e. they define access points. There are 99 values defined in the Bib-1 attribute set. Among the Use values, values are corresponding to the author, title, keywords, year of publication, etc.

Among the Use attributes described in Table 2, some are associated with several fields simultaneously (“Author-name-and-title”), and the value “Anywhere” is associated with all search fields at once.

Table 2

Use Bib-1 attributes

Attribute number	Attribute name
4	Title
21	Subject
31	Date-publication
1003	Author
1000	Author-name-and-title
1035	Anywhere

Among the Use attributes, some are associated with several fields at the same time. The last value is associated with all search fields.

Attributes 2: Relation (Relationship attributes).

The Relation attributes are described in Table 3, the relationship between access points and search terms, i. e. they indicate how the search term relates to the selected data from the fields defined by the Use attribute.

Table 3

Relation attributes of Bib-1

Attribute number	Attribute name
1	Less than
3	Equal
6	Not equal

Attributes 3: Position.

The Position attributes are described in Table 4, the attribute defines the position of the search term inside the field or subfield, i.e. they indicate where the search term should be located in the field defined by the Use attribute.

Table 4

Position attributes of Bib-1

Attribute number	Attribute name
1	First in the field
3	Any position in the field

Attributes 4: Structure (Structure attributes).

The Structure attributes indicate which structure the search term has (Table 5).

Attributes 5: Truncation (Truncation attributes).

The Truncation attributes indicate whether one or more characters located in the position defined by the Truncation attributes can be ignored when matching with the search term (Table 6).

Table 5

Structure attributes of Bib-1

Attribute number	Attribute name
1	Phrase
2	Word
4	Year

Table 6

Truncation attributes of Bib-1

Attribute number	Attribute name
1	Right truncation
100	Do not truncate
102	RegExpr-1

Attributes 6: Completeness (Completeness attributes).

The Completeness attributes indicate what is a mandatory area of matching when searching, i.e. whether the search term represents a complete, incomplete subfield or field, or in other words, whether additional words can be present together with the search term in the field or subfield (Table 7).

Table 7

Completeness attributes of Bib-1

Attribute number	Attribute name
2	Complete subfield
3	Complete field

Thus, to search for records in which the author Ivanov or Ivanova occurs, you need to set:

bib-1 1=1003 2=3 3=3 4=2 5=1 6=2 "Ivanov".

Types of requests in Z39.50

There are several types of requests provided in Z39.50:

-- Query Definitions

```

BEGIN
EXPORTS Query;
Query ::= CHOICE{
  type-0 [0] ANY, -- Any
  type-1 [1] IMPLICIT RPNQuery, --RPN
  type-2 [2] OCTET STRING, -- CCL ISO 8777
  type-100 [100] OCTET STRING, -- CCL Z39.58
  type-101 [101] IMPLICIT RPNQuery, -- RPN
  type-102 [102] OCTET STRING,
  type-104 [104] IMPLICIT EXTERNAL}-- SQL
END
    
```

In this case, a type-0 request is any query in the syntax of the DBMS that the server is associated with. The target must pass requests of this type to the database provider without modification.

Type-2 and type-100 queries are queries in CCL syntax. They are rarely used in Z39.50 and will not be discussed here.

Type-104 – SQL queries are of interest. This is a new type of request that has been included in the Z39.50 standards since February 2000. Today, there are practically no servers that support SQL queries in Z39.50. Nevertheless, its definition should be given:

```

--SQL Query Definition
SQLQuery ::= SEQUENCE {
  abstractDatabaseFlag [0] BOOLEAN OPTIONAL,
  queryExpression [1] IMPLICIT InternationalString}
-- as defined in the SQL Standard [ISO/IEC9075]
-- with/without Z39.50 schema abstraction extension as
-- specified by the flag
-- end SQL Query Definition
    
```

As you can see, the SQL query is a simple text string. However, it can be built in two ways: in the usual way (abstractDatabaseFlag=FALSE) and through an abstract data schema (abstractDatabaseFlag=TRUE). In the following chapters, the discussion of this type of request will continue.

The most interesting queries are type-1 and type-101-RPN requests (RPN – Reverse Polish Notation). For version 3 of the Z39.50 protocol, both types are no different. Type-1 requests (RPN) are mandatory for all Z39.50 servers. Support for other types of requests by the Z39.50 servers is optional.

**4.4. Distinguishing features of the Z39.50 protocol and the basic idea of presenting information**

At the moment, the protocol has received a broader scope of application. Today, with the help of Z39.50 technology, it is possible to access scientific and technical, biological, museum data, reference information, and so on.

Currently, the Z39.50 Support Agency is actively working to create a new version of the protocol, which will have a significant difference from previous versions, consisting in supporting SQL as one and valid query languages.

Initially, the HTTP protocol was used as the standard for accessing distributed databases. But the HTTP protocol does not allow unifying access to heterogeneous information. This problem can be solved only with the help of auxiliary tools – programming languages.

Therefore, information resources that do not support the Z39.50 protocol are isolated and heterogeneous, which complicates working with them. The user, working with only one special application built according to the Z39.50 protocol, can search through various remote databases.

The Z39.50 protocol provides for the existence of various sets of attributes for searching and syntactically describing records. The Z39.50 Maintenance Agency maintains a register of attribute sets [18].

The abstract database represented by the protocol displays a specific model of an existing database. The developer will have to correctly form the structure of a real database so that an abstract model can be created on its basis.

Distinguishing features of the Z39.50 protocol:

1. The information presentation model embedded in the protocol does not depend in any way on the information sources using this protocol. In other words, the protocol provides a kind of abstract model for presenting information at each stage of client-server interaction.

2. The Z39.50 protocol fully provides session interaction between the client and the server. This feature is embedded in the protocol itself and is implemented in all its applications, whether it is a server system or a client program.

The basic idea of presenting information when working with the Z39.50 protocol lies in abstracting from the specific

structure of any database. To do this, the standard describes a kind of abstract database model. This model includes a complete set of elements necessary for accessing and processing information stored in the database. The abstract model describes in the form of separate elements not only, for example, possible search fields or information output formats, but also all operations performed by the server.

Each element of this abstract model is described in detail to an unambiguous interpretation and standardized with the assignment of a unique identifier – OID. Work with each specific DBMS should be organized only through this abstract model by exchanging data packets (APDUs) containing sequences of objects identified by names (labels).

## 5. Results of use based on Z39.50 protocols

### 5.1. Exploring the capabilities of the Z39.50 protocol

The Z39.50 protocol describes the network interaction of subjects in the client-server architecture. However, this interaction is somewhat different from the classic client-server architecture, in which only the client can be the initiator of any request, and the server is always assigned the passive role of waiting and responding. As will be seen below, this is not always the case in Z39.50. Maybe that's why the protocol developers changed the terminology, replacing the terms "client" and "server" with the terms "origin" and "target", respectively. With rare exceptions, the concepts of "client" – "origin" and "server" – "target" coincide.

To describe the logic of network interaction in Z39.50, the standard defines the following components:

- Origin – the component that initiates the Z39.50 communication session;
- Target – the component waiting for the Z39.50 session to open;
- Database user – data user;
- Database provider – data provider;
- Service provider – a component that accepts network requests and responses;
- Service user – a component that uses network requests and responses;
- Client – an application that includes origin and database user;
- Server – an application that includes target and database provider.

The standard describes four service primitives, in terms of which the rules of service procedures Z39.50 are formulated:

- Request (request) – a primitive used by origin to initialize its service provider of the corresponding service procedure;
- Indication – a primitive that transmits information to target from the service provider to the service consumer;
- Response – a primitive used by the target to initialize its response service provider.
- Confirmation (notification) – a primitive passed to origin from the service provider to the service consumer.

The sequence of the service procedure can be illustrated by the example of a search query initiated by origin:

- origin – the service consumer forms a SearchRequest request for its service provider;
- a special package is sent from origin to target via the network – APDU SearchRequest;

- target – the service provider indicates to its service consumer about the receipt of the SearchRequest;
- target service – the consumer generates a SearchResponse for his service provider;
- a special package is sent from target to origin – APDU SearchResponse;
- origin – the service provider notifies its service consumer about the SearchResponse.

This is usually what all service procedures look like (Init, Search, Present, Delete, Resource-report, Sort, Scan, Extended-services). However, there are also those in which the roles of target and origin are reversed. The initiator of the procedures (Access-control and Resource-control) is a target.

### 5.2. Organization of requests using the Z39.50 protocol

According to the Z39.50 protocol, search queries are formulated not to a real database, but an abstract one. In other words, the data is extracted not from the database directly, but from intermediate sets created by the server at the time of the request. Intermediate data sets are characterized by search attributes that are used to compose a search query.

The Z39.50 protocol supports a mandatory type of request in reverse Polish notation, which is called an RPN request (Reverse Polish Notation). This query can have a complex structure that contains a combination of attributes and search terms. Attributes are characterized by a set of parameters that define the search rules for each term. The request can be displayed as a string for clarity. To specify a search attribute, you need to write a combination of two numbers (the first of which is the type, and the second is the value). Thus, each field of the database table can be represented as an abstract record.

The thesaurus database is searched using a fixed set of attributes (Bib-1, XD-1, util, and Zthes-1), which are included in the Use attribute group (type 1). Also, five types of additional attributes are used to build a query (Relation (type 2), Position (type 3), Structure (type 4), Truncation (type 5), Completeness (type 6)) that define the query. The most common set of attributes is Bib-1, which includes search attributes such as Author, Title, DatePublication, etc. in the Use type.

An example of an RPN query expanded into a string (a query in the PQF – Prefix Query Format syntax) can be a request to search for records in which the author begins with "Serikbayeva" and occurs at any position of the field:

```
@attrset Bib-1 @attr 1=1003 @attr 2=3 @attr 3=3 @attr
5=1 {Serikbayeva}
where
@attrset Bib-1 – setting a set of attributes;
@attr 1=1003 – corresponds to the Author field,
@attr 2=3 – equal,
@attr 3=3 – any position in the field,
@attr 5=1 – right truncation,
Serikbayeva – search term.
```

### 5.3. Building Reverse Polish Notation Requests

The RPN request can be represented as a tree, in the nodes of which there are binding operators (AND, OR, AND-NOT). The leaves of this tree are the "term attributes" (APT) blocks. Fig. 4 schematically shows the RPN request.

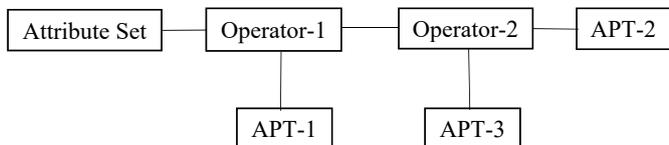


Fig. 4. The structure of the reverse polish notation request

Each attribute is a pair that reflects the attribute type and its value. The following search attributes are present in the given example: type “Use” and value “Creator-Name” (1=2035), type “Relation” and value “Equal” (2=3), type “Position” and value “Any position in field” (3=3), type “Structure” and value “Word” (4=2), type “Truncation” and value “Right truncation” (5=1) and type “Completeness” and value “Incomplete subfield” (6=2).

As a result of performing a database search, the client can receive the following information from the server: an error message, the number of records found, or the records found themselves. The first answer option is associated with an error, the second and third options correspond to a successful search. Which of them will be received by the client depends on the parameters that are transmitted to the server along with the search query.

For this purpose, the concepts of small, medium, and large sets are introduced. Here, a set is understood as a set of found records, numbered end-to-end. All records from the small set are always returned, all records from the large set are never returned, and some records are returned from the medium setting. Next, the following parameters are set:

- the upper limit of the small set, i.e. the maximum number of records in the small set, which begins with the first record;
- the lower bound of the large set, i. e. the number starting from which the records fall into the large set. All records whose numbers are greater than the upper limit of the small set, but less than the lower limit of the large set, are considered records from the middle set;
- the number of returned records from the average set.

By changing these three parameters, you can return any number of records, including none.

Finally, it should be noted that the server should save all the records found during the search in the session block for later use. If the server allows the option to assign a name to the search result, this saved population can be assigned a name, if not, the population is kept unnamed and rewritten during the subsequent search. Named result sets that are stored on the server can be used in subsequent RPN requests, where they act as the same operands as APT blocks.

In addition to what is said in the figure, a set of attributes is specified, which is used by default.

In Z39.50, the RPN request is not a string of characters but is a structure that can be represented as a string but only for clarity.

As a result of performing a database search, the client can receive the following information from the server in the APDU SearchResponse:

- error message;

- number of records found;
- the records found themselves.

The first answer option is associated with some error that is diagnosed, for example, by bib-1. In particular, you can get the message 236 Access to specified database denied (Access to the specified database is prohibited).

Creating efficient and adaptive distributed systems allows you to significantly speed up data processing. To consider this issue, we will analyze the problems that arise during the design and operation of distributed systems.

For distributed information systems that include many different databases with different structures and content, the issue of searching for information in databases using ontologies, thesaurus, and classification schemes presented in the form of separate databases is very relevant.

There are many different ways to build databases, organize access to their contents, and implement explicit and implicit links between the database and other information resources. Many of these methods are based on strict ontological models and, for practical implementation, impose very strict requirements on the organization of information systems and databases, up to the complete overload of information into intermediate storages, the functional properties of which make it possible to identify all semantic relationships between information objects based on specified ontological models. Such an approach has a right to exist, but the question remains how to enable the search for semantically related information in existing distributed information resources, and in the case when they cannot be overloaded into specialized repositories [8].

As a result of using this protocol, it is possible to create distributed information systems that include databases of various organizations.

In a distributed environment, data synchronization mechanisms should be involved, for example, based on replication. At the same time, standard protocols should act as network communication protocols, for example, OAI, Z39. 50, SRW/SRU, LDAP, etc. (Fig. 5).

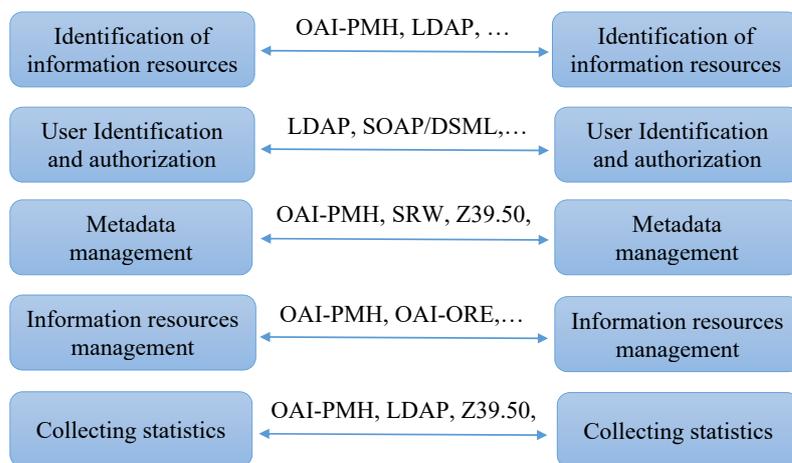


Fig. 5. Network interaction of subsystems of distributed information systems

The practical implementation of SRW/SRU services will give a significantly new quality of the information system – the ability to include its resources in global search engines at a higher level than the level of external indexing of static web pages by other systems. Other possible types of search are related to the search for the specified templates and the search involving the ontology. The latter is a more intelligent

type of search. Its implementation requires additional information about the subject area, including definitions of terms, entities, and relationships.

Currently, there are quite powerful information systems that meet the needs of researchers in information to one degree or another. However, the main drawback of most systems is the limited possibilities of ensuring the integration of resources both inside each of the systems and outside. It should be noted that the basis for the development of IP is, first of all, standards and international recommendations that form the IP profile. It is understood as a set of one or more basic normative and technical documents (standards and specifications) focused on solving a specific task (implementation of a given function or group of functions of an application or environment), indicating, if necessary, selected classes, subsets, options of basic standards required to perform a specific function. The most important is the metadata profiles of the information circulating in the system. The choice of a profile should be based on the following requirements:

- include the main types of information required to support scientific work;
- be open, i. e. provide access to relevant information on these descriptions;
- be extensible, i. e. provide the possibility of detailing descriptions;
- provide information integration capabilities;
- provide opportunities for unique identification of information;
- provide the ability to host and search for information in a distributed environment;
- be focused on modern and promising technologies for describing and using information;
- provide opportunities for interoperability with the external environment.

The implementation of each subsystem with standardized external interfaces is not very significant. However, the basic technologies of their implementation follow naturally from their general functionality (Fig. 6).

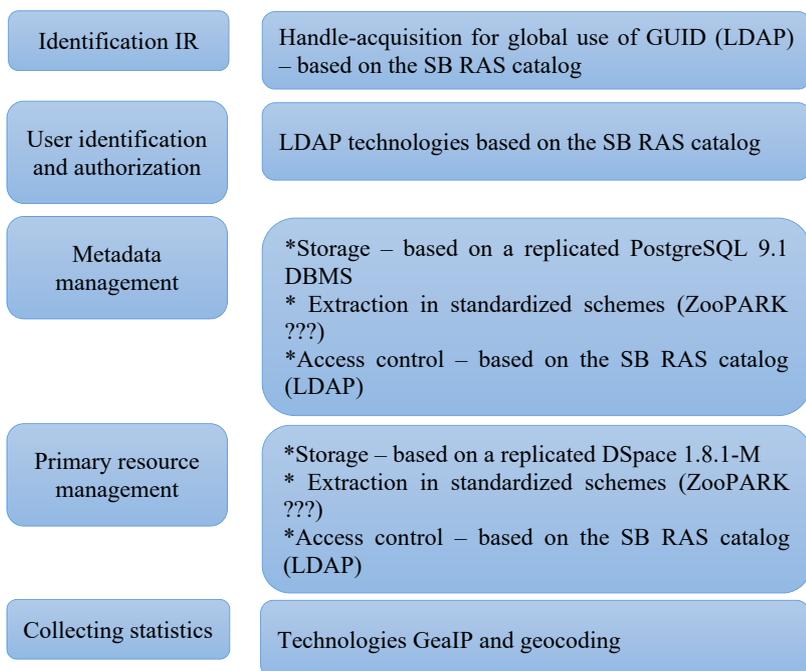


Fig. 6. Basic technologies of the distributed information systems component

The developed model of the information system can be used as a standard model of the system for working with documents related to scientific and educational activities since it solves the main tasks imposed on these systems: providing a system of reliable long-term storage of digital (electronic) documents while preserving all the semantic and functional characteristics of the source documents; providing “transparent” search and user access to documents, both for familiarization and for analyzing the facts contained in them; organization of information collection on remote digital repositories that support the OAI-PMH, SRW/SRU, Z39.50 protocols [11].

#### 5. 4. Application for the transformation of abstract requests

Because of the work, a function was created in the built-in PostgreSQL DBMS language, which performs the transformation of an abstract RPN query into a real SQL query to the thesaurus database. Below is an example of accessing a function in the program:

```
$query = pg_query(“SELECT convert(“.$qStr.”)”).
```

Let’s move on to the algorithm of the function. The input of the function is a string containing an abstract RPN request. The function reads the string of an abstract RPN request, then determines the number of subqueries in the request and the relationship of subqueries to each other. Next, the fragments of the SQL query are substituted instead of the search attributes. In the end, parentheses and logical operators are arranged if the abstract RPN request has a complex structure. The function, in turn, returns a string with the generated SQL query at the output.

PostgreSQL functions were used to implement the algorithm for transforming an abstract RPN query into an SQL query. The PostgreSQL DBMS supports search by phrase (a set of words taking into account the order), a list of words, and a set of characters [9].

For convenient search through the thesaurus database, an interface was developed in PHP, using which the user can form an abstract query to the thesaurus database.

To access the thesaurus database, the user should build an abstract query. For this purpose, a custom WEB application was developed (see Appendix B), which generates abstract queries to the thesaurus database [12]. To get an abstract query, the user must fill in the form input fields with the following search parameters: the name of the search term, the name of the set of attributes, a search attribute of the Use (Access Point) type. If desired, the user can build queries to search for more accurate information using additional search attributes and logical operators. Fig. 7 shows the parameters of the abstract query and ways to search for terms.

After filling in the data, you must click on the “Add” button. The selected search parameters will be displayed in the input field below. Fig. 8 shows the parameters for searching for additional attributes.

**Поисковый терм:**  
 Информационная система

**Имя набора атрибутов:**  
 XD-1

**Поисковое поле:**  
 1-Title

**Дополнительный атрибут:**  
 Выбрать дополнительный атрибут

**Логический оператор:**  
 Выбор логического оператора

Добавить    Следующий уровень

**Строка запроса:**

Fig. 7. Choice of parameters of an abstract query

**Дополнительный атрибут:**  
 Выбрать дополнительный атрибут

**Логический оператор:**  
 Выбор логического оператора

Добавить    Следующий уровень

**Строка запроса:**  
 @attr XD-1 1=1 Информационная система

Очистить    Сформировать запрос

Fig. 8. Choice of parameters of an abstract query

№	title	link_id	term_qualifier	term_vocabulary	description	document_language	term_category
7	Информатика	54F38E0C	publ555		Информатика — это наука, которая занимается вычислением, хранением и обработкой информации. Она развивается вместе с компьютерами и сетью интернет, а потому базируется на компьютерной технике и невозможна без нее.	ru	
2	Информационная система	BFC88BB8	abacus_ru	-	Информационная система — это взаимосвязанная совокупность средств, методов и персонала, используемых для хранения, обработки и выдачи информации для достижения цели управления.	ru	-

Fig. 9. Search result

To transform an abstract query into a real SQL query to the thesaurus database, a function was developed that is embedded in the DBMS, which works according to the following algorithm.

To begin with, the user must build an abstract query, for example, using the application and check its correctness using the application:

@or @attr XD-1 1=1 Information system @attr XD-1 1=1 @attr 5=1 Informatics

This request in the form of a string is passed to the function as an input parameter. Next, the request is parsed:

@or – logical OR operator  
 @attr XD-1 1=1 – search by “Term name”

Information system – search term  
 @attr XD-1 1=1 – search by “Term name”

@attr 5=1 – right truncation  
 Informatics – search term

The result is replaced with fragments of the SQL query:

@or -> or  
 @attr XD-1 1=1 -> title  
 Information system -> “Information system”

@attr XD-1 1=1 -> search by title  
 @attr 5=1 -> LIKE  
 Informatics -> “Informatics%”

Next comes the arrangement of logical operators and brackets:

((title = “Information System”) or (title LIKE “Informatics%”))

After that, an SQL query is executed against the thesaurus database:

SELECT \* FROM zthes\_cat WHERE ((title = “Information System”) or (title LIKE “Informatics%”))

Fig. 9 shows the result in the form of tables based on the selected attributes.

All application parameters (including attribute sets) are external to the application, so it can work with any thesaurus database (you only need to replace many parameters).

## 6. Discussion of the results of the development of queries using the Z39.50 protocol in distributed information systems

The paper describes the development of an application for constructing abstract RPN (PDF) queries of the Z39.50 protocol to the thesaurus (from query generation to its execution), and also considers an alternative approach based on the CQL query language.

Using the developed application solves the problem of unified access to information. Since there are a large number of thesauri in the global network, each of which has its own form of presentation and storage of information. Each database has a unique information storage structure, where each field has its own name and purpose.

The developed WEB application interactively creates an RPN request, checks its correctness and executes it. Moreover, the first two tasks are solved on the client's machine. Queries can be simple and complex. The whole range of queries from the Z39.50 archive has been implemented, including searching for "phrases" and "character sets" by some search term.

Abstract queries are formed based on search attributes from the sets Zthes-1, bib-1, XD-1, and util. These attribute sets are included in the Use attribute group [10] and are intended for thesaurus search.

Simple queries are constructed without using logical operators and perform a search by a single search parameter. Complex queries include logical operators that link several simple queries together. To get a more accurate search result, five groups of additional attributes are used when building a query (Relation, Position, Structure, Truncation, Completeness). Queries can be simple and complex. A recorded WEB application that interactively creates an RPN request, verifies its correctness, and executes it.

Z39.50 is a client-server architecture standard in which the search engine and the interface are divided into independent parts. If both the client and the server meet the standard, the Z39.50 client can search for any brand of the Z39.50 server. The most common databases in different local systems can be found through the same local client or interface. This does not solve the problem of how the interface should look or how it should act, it is up to the user to choose the interface. The connection of library systems with

the Internet and the development of the Z39.50 protocol open up the prospect of access to an ever-growing array of bibliographic databases and full-text databases through a local automated system. The ability to directly connect users to resources providing various computing platforms has increased the attractiveness of the Z39.50 protocol for libraries linking institutional systems. As a result of using this protocol, it is possible to create distributed information systems containing databases of various organizations.

Support of information systems in the field of scientific and educational activities is relevant, since the need for information always exists. In order to satisfy this need, it is necessary to organize access to various resources.

In the future, it is planned to supplement this work.

## 7. Conclusions

1. Ways to use the Z39.50 protocol in the creation of distributed information systems are considered and described. a system for exchanging information using the Z39.50 protocol using elements of the main attributes of the protocol has been developed.

2. Ways to generate requests using the Z39.50 protocol have been considered and effective solutions have been made.

3. The Z39.50 protocol provides a full cycle of work with abstract RPN (PQF) queries (from query execution to execution), and provides an alternative approach based on the SQL query language. Queries can be simple or complex. Protocols for exchanging information between servers using the Z39.50 protocol were demonstrated.

4. Using the Z39.50 protocol, abstract queries are organized. Abstract requests are formed based on search attributes from zthes-1, bib-1, XD-1, and util. These sets of attributes are included in the Use attributes group and are intended for a search for the thesaurus.

## Acknowledgments

"This research has been funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP09057872)"

## References

1. Temirbekov, N. M., Zhizhimov, O. L., Baygereev, D. R., Omirzhanova, B. B., Temirbekov, A. N., Orazgalieva, A. M. (2018). Postroenie nauchno-obrazovatel'nogo klastera na baze integrirovannoy raspredelennoy informatsionnoy sistemy Akademgorodka. CITech-2018. Ust'-Kamenogorsk, 119–126. Available at: [http://www.ict.nsc.ru/jspui/bitstream/ICT/1879/10/Part1\\_119-126.pdf](http://www.ict.nsc.ru/jspui/bitstream/ICT/1879/10/Part1_119-126.pdf)
2. Shokin, Y. I., Temirbekov, N. M., Zhizhimov, O. L., Temirbekov, A. N., Baigereyev, D. R. (2018). Model of the integrated distributed library information system of Almaty Akademgorodok. Computational technologies, 23 (5), 110–119. doi: <https://doi.org/10.25743/ict.2018.23.5.009>
3. Sadirmekova, Zh. B., Zhizhimov, O. L., Tusupov, D. A., Sambetbaeva, M. A., Murzakhmetov, A. N. (2019). Requirements for information systems to support scientific and educational activities. Trudy XVII Mezhdunarodnoy konferentsi DICR-2019. Novosibirsk, 171–177. doi: <https://doi.org/10.25743/ICT.2019.77.63.026>
4. Serikbayeva, S. K., Tussupov, J. A., Sambetbayeva, M. A., Yerimbetova, A. S., Taszhurekova, Zh. K., Borankulova, G. S. (2021). EduDIS construction technology based on Z39.50 protocol. Journal of Theoretical and Applied Information Technology, 99 (10), 2224–2255. Available at: <http://www.jatit.org/volumes/Vol99No10/6Vol99No10.pdf>
5. Kurmanbekovna, S. S., Gabitovich, B. A., Aralbaevna, S. M., Bakirbaevna, S. Z., Sembekovna, Y. A. (2021). Development of technology to support large information storage and organization of reduced user access to this information. International Journal of Advanced Computer Science and Applications, 12 (7), 493–503. doi: <https://doi.org/10.14569/ijacsa.2021.0120757>

6. Serikbayeva, S. K., Tussupov, J. A., Sambetbayeva, M. A., Serikbayeva, J. K. (2021). Requirements for distributed information systems to support scientific and educational activities. *Bulletin of the South Ural State University. Ser. Computer Technologies, Automatic Control & Radioelectronics*, 21 (3), 149–160. doi: <https://doi.org/10.14529/ctcr210315>
7. Shokin, Y. I., Fedotov, A. M., Zhizhimov, O. L. (2015). Technologies for designing of distributed information systems to support research. *Vychislitel'nye tekhnologii*, 20 (5), 251–274. Available at: <https://cyberleninka.ru/article/n/tehnologii-sozdaniya-raspredelennyh-informatsionnyh-sistem-dlya-podderzhki-nauchnyh-issledovaniy>
8. Zhyzhymov, O. L., Fedotov, A. M., Fedotova, O. A. (2012). Building a generic model of information system for working with documents on the scientific heritage. *Vestnik Novosibirskogo gosudarstvennogo universiteta. Seriya: Informatsionnye tekhnologii*, 10 (3), 5–14. Available at: <https://nsu.ru/xmlui/bitstream/handle/nsu/236/01.pdf?sequence=1&isAllowed=y>
9. Mazov, N. A., Zhizhimov, O. L., Skibin, S. V. Experience in Designing Distributed Information Systems on the basis of Z39.50 Protocol. Available at: <https://www.gpntb.ru/win/inter-events/crimea2002/trud/sec2/Doc4.HTML>
10. Fedotov, A. M., Barakhnin, V. B., Zhizhimov, O. L., Fedotova, O. A. (2014). A model of information system to support scientific and educational activities. *Vestnik Novosibirskogo gosudarstvennogo universiteta. Seriya Informatsionnyy tekhnologii*, 12 (1), 89–101. Available at: <https://cyberleninka.ru/article/n/model-informatsionnoy-sistemy-dlya-podderzhki-nauchno-pedagogicheskoy-deyatelnosti>
11. Bayekeyeva, A., Tazhibayeva, S., Beisenova, Z., Shaheen, A., Bayekeyeva, A. (2021). Controlled Multilingual Thesauri for Kazakh Industry-Specific Terms. *Social Inclusion*, 9 (1), 35–44. doi: <https://doi.org/10.17645/si.v9i1.3527>