Development of Technology to Support Large Information Storage and Organization of Reduced User Access to this Information

Serikbayeva Sandugash Kurmanbekovna[1]
Sadirmekova Zhana Bakirbaevna[4]
Department of Information Systems
L.N. Gumilyov Eurasian National University
Nur-Sultan, Kazakhstan

Sambetbayeva Madina Aralbaevna[3]
L.N. Gumilyov Eurasian National University
Institute of Information and Computational Technologies
CSMES RK
Nur-Sultan, Almaty, Kazakhstan

Batyrkhanov Ardak Gabitovich[2]
Department of Information Systems
L.N. Gumilyov Eurasian National University
Atyrau State University Named after H. Dosmukhamedov
Nur-Sultan, Atyrau, Kazakhstan

Yerimbetova Aigerim Sembekovna[5]
Institute of Information and Computational Technologies
CSMES RK
Almaty, Kazakhstan

***Abstract*—This article solves the problem of developing a technology for supporting large information storages and organizing delimited user access to this information, which provides a service both for 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 relationships between them. It will also allow the development of 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. In existing DL developments, as a rule, search and access to information are provided only through visual graphical interfaces. The task of the subsystem for integrating various digital resources is to provide other subsystems with a single interface for access to information stored in the data sources of the system. That is, any resource must 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.**

Keywords*—*Information systems; digital library; metadata; collection; privilege; rights; administrator

I. Introduction

Information systems (IS) to support scientific and educational activities operate with various kinds of information, such as publications, digital documents, electronic collections, ontological descriptions, data arrays, logical descriptions, etc. As a rule, these resources, which are in demand by different groups of researchers, are inaccessible due to the impossibility of their search and identification. Semantic links between information resources increase their value and provide additional opportunities for information retrieval and identification. Data integrated into open semantic space is a body of knowledge about a certain subject area in the form of a semantic structure, based on which qualitatively new scientometric measurements and studies of the structural properties of the body of scientific knowledge become possible.

The main purpose of the study is to develop a conceptual vision technology of using the digital library management system as a typical system for storing and accessing information resources and a description of the basic requirements for its implementation and operation.

Information storage technology is a kind of integrated technology designed to implement procedures, methods, and means of storing and using a database complex in solving user problems. Large amounts of data can be stored on one or more servers. These arrays are usually called information storage [1].

Data presentation; As a subject area, the materials of the scientific heritage in the field of IT technologies were considered. In the information space, events, facts, and any other entities of the real world exist only in the form of documents. The document is the main object involved in any information system (IS). The main function of the document is informational, i.e. the ability to satisfy the information needs of a person. Many documents containing factual information, having the same physical structure and logical, informative purpose, form collections. Collections, according to [2], are characterized by their descriptions and descriptions of the structure of documents from which it consists, and represent a systematized collection of documents, united by some criterion of belonging, for example, by content, purpose, access method, etc. provided with meta description (metadata) under standards and data schemas. Collections can be nested

within each other, but one collection can only have one parent collection. Any document can be placed in several collections.

Metadata is created in the process of explicit or implicit cataloging and corresponds to one or another generally accepted area. To extract metadata, algorithms are used for parsing the title page, extracting keywords in their absence, an algorithm for constructing an abstract, based on a graph connected to the text. The main catalog of information resources of the IS metadata server is built under the metadata schema. The developed schema takes into account the basic requirements of the Dublin Core metadata schema. In what follows, our metadata schema will be called internal.

Data storage: For long-term storage of documents, the institutional repository Dspace [3] was used, due to its ability to expand the list of supported metadata, which allows it to be customized for different subject areas. To support the process of populating full-text databases, workflows and user interfaces were configured: the created metadata profiles were registered in the DSpace system.

Data exchange: To implement the exchange of metadata between DSpace, under the extended profile, a service in XSLT was created that converts metadata schemas from the internal DSpace schema to the metadata server schema and the Dublin Core schema using qualifiers. The OAI service is also implemented, which periodically, in batch mode, under the schedule, synchronizes the metadata of the repository and the metadata server. To fill the main catalog of metadata under the created metadata schemas, controlled vocabularies from the reference block of the system are used [4, 27].

Environment functioning: Based on the use of Z39.50 and LDAP protocols. At the same time, mechanisms are provided for converting data from subject schemas to the abstract schema of the Z39.50 protocol. The virtual environment consists of a registry of objects and resources, the main Z39.50 servers, several functional modules, and a web interface with public and administrative sections for accessing various functions of the environment. For each source, a separate Z39.50 servers is installed, which transforms data from the source schema into an abstract data schema [5, 28].

Collecting data from external sources: The chosen technology of integration of the developed information system to support research on scientific heritage with a DD (digital depository) allows using any other DD implementations that support the OAI protocol. To work with external data centers that support the OAI protocol, service has been implemented that converts the Dublin Core metadata schema into the system's internal metadata schema.

Metadata extracted from external data centers is also placed in the database of the digital library (metadata server). If this is possible for a specific DD, then the metadata is extracted in an XML schema with further transformation into a GOST schema. If the data schema of the external data center is unknown, then the metadata is converted from the Dublin Core-based schema to the GOST metadata schema.

To integrate the digital library with external systems and applications using the OAI protocol, transformation services of the internal metadata schema are used, for example, in MARC XML 20 schema or RUSMARC, etc. If necessary, the same approach can be used to transform metadata into other schemas of the MARC family.

Search for data: The functionality of searching for documents (information resources) is available to end-users in three ways: through the user interface of the information system (metadata server), through a specialized search service (for external applications), and the user interface of the DSpace DD [3]. However, in any case, the documents themselves are always stored in the DD, therefore, through whatever interface the user finds the document he needs, the document will be directly retrieved via the HTTP protocol from the DD.

User access control: At all levels of user interfaces (metadata server, services, and DD), user access control to the resources of the information system is carried out based on identification information under the LDAP protocol. DSpace has built-in support for LDAP, and in programmed services and the metadata server, support is provided by the services of the information system.

The proposed technology for supporting large information storages is based on the client-server architecture of the IS and meets the requirements described in [4]: {{1}}.

- a unification of the process of exchange of scientific research results;
- operating with data and documents integrated into an open semantic space;
- provision of services for transforming heterogeneous resources that implement means of description, presentation, automatic linking of resources, as well as interaction with search and classification mechanisms by the needs of users.

II. Management System for Digital Libraries

A. Administrative Collections

The information system consists of objects - elementary units of documents, from documents - information units. Many documents containing factual information, having the same physical structure and logical, informative purpose, form collections. Collections are characterized by their descriptions and descriptions of the structure of documents from which it consists.

A collection is a set of documents united by semantic attributes and having the same structure.

A collection is a set of documents with a dedicated fixed structure, the content of which has the same thematic focus. From the point of view of unification of work with documents, we will represent the information system as a set of collections [6].

A collection is a common form of organization of information resources, which is determined by its parameters (style, attributes) and the structure of the documents included in it and is a systematized set of documents, united by some criterion of belonging, for example, by content, purpose, access method, etc., provided with meta description

(metadata) following data standards and schemas. The document is characterized by its parameters (style, attributes) and the structure of the objects of which it consists. An object is determined by the type of data (following the selected data schema) that it contains, the description of the properties and methods of the object.

Metadata Server digital library management system (DLMS) contains an administrative (service) collection of the Main Catalog contains descriptions of the QDC metadata schema, extended with metadata and for MECOF compliance and descriptions of service metadata describing the structure of objects, user interfaces, associative links between documents, access lists to documents to differentiate access rights to the object by the rights of the actors (if desired, it can be extended with new metadata) [7].

Fig. 1 is represented by administrative collections: users (actors), sets of metadata schemas and access lists to documents (collections).

A metadata schema is a set of metadata elements used, for example, to describe specific types of information resources. A set of metadata elements are defined by the DL super administrator that defines various schemes, from relatively simple text materials to more complex multimedia objects and DL collections, at the earliest stages of library creation, taking into account the needs of all potential users. However, over time, it can be expanded and the metadata schema will be adjusted to meet new requirements.

The system operates with the following types of information objects - collections, documents, objects.

From a functional point of view, information resources of a collection are divided into data (document) and metadata. Metadata is specially prepared, machine-readable, structured information about a resource, representing the properties that the resource has, the services that the resource provides.

Collection metadata is a set of values of some of its properties and properties of information resources belonging to it. The specific functions of metadata and their composition can vary significantly, depending on the specifics of DL and a particular collection. In our case, the set of metadata is divided into the following classes, such as: unique (identifier (UID)), automatic (owner, record modification date), public (informational content of the record), and service (administrative metadata, access rights, user data, etc.) metadata. Accordingly, each of which is also subdivided according to the level of access: automatic, administrative, and public [8].

Fig. 1. Administrative Collection.

The metadata system is the central logical component of any collection of documents. Collection metadata describes the properties of documents, sections, and the collection as a whole. Metadata describes the structure of the collection, determines the composition of the collection, and ensures correct interpretation and processing of the documents presented in it. Collection metadata also describes the structural properties of documents (types, relationships), their presentation formats, access control, resource content, information about authors, the collection's classification system, etc.

Thus, the main element of the considered scheme is structural metadata (that is, formalized descriptions of the structure of stored information). It makes it easy to build a library with a distributed data warehouse: the connection between the warehouse and the system will provide meta-descriptions of the warehouse data.

The digital library management system is an information system that provides the development and administration of digital library systems, as well as the integration of software that offers advanced and specialized additional features.

Fig. 2 shows a scheme for managing collections and documents in the electronic library management system.

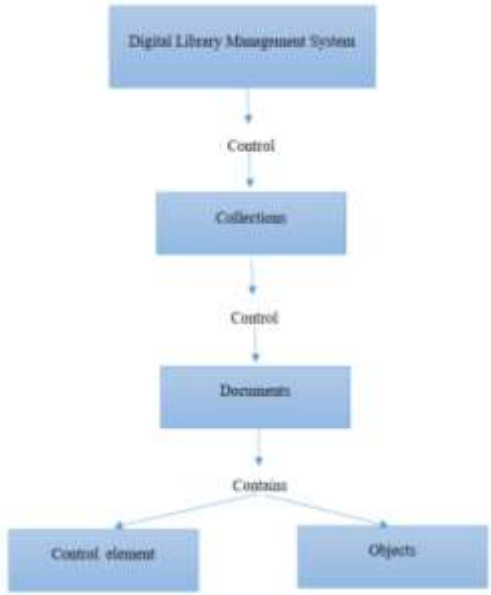

Fig. 2. Digital Library Management Systems.

Main functions of the Digital Library Management System:

- Access to resources-navigation through DL collections, query processing, location detection, extraction, resource transformation.
- Resource management – creating new information objects and collections, adding, deleting, etc.
- Metadata and dictionary management-creation, up-to-date maintenance and development, the composition of metadata and dictionaries is determined by agreements, taking into account international standards.

- User management – their registration, registration of rights, personal information.
- System administration – installation, configuration, recovery, maintenance, data security.

B. Basic Concepts in Access Control

Identification-assigning identifiers to users (unique names or labels) under which the system "knows" the user. In addition to user identification, user groups, system resources, and so on can be identified. Identification is also needed for other system tasks, such as event logging. In most cases, identification is accompanied by authentication [9, 10].

Authentication - authentication-verification of the identity of the user presented to them. For example, at the beginning of a session in the system, the user enters a name and password. Based on this data, the system performs identification (by user name) and authentication (by matching the user name and the entered password).

Access control is a method of protecting information by regulating the use of all system resources.

The identification and authentication system is one of the key elements of the infrastructure of protection against unauthorized access of any information system.

The access model operates with the terms "entity", "subject", "object".

Entity – any named component of the system being developed.

A subject is an active entity that can initiate resource requests and use them to perform any operations. The subject can be a program running in the system or a "user" (not a real person, but the essence of the system).

An object is a passive entity used to store or retrieve information. An object can be considered, for example, a collection, a document, etc.

Access – interaction between a subject and an object, as a result of which information is transferred between them.

Two fundamental types of access: a read-an operation that results in the transfer of information from the object to the subject; a write-an operation that results in the transfer of information from the subject to the object.

The basis of the mandatory security policy of the system is the mandatory access control, which implies that:

- all subjects and objects must be identified;
- a hierarchy of information security levels is set;
- each object of the system is assigned a level of protection that determines the value of the information contained in it – its level of availability;
- each subject of the system is assigned a restriction level, which determines the level of trust in it – its access level; and
- the decision to allow the subject to access the object is made based on the type of access and the comparison of the hierarchy of the subject and the object.

The main purpose of the mandate model is to prevent access to resources of objects with a high level of access to objects with low-level access.

In the model, objects and subjects are categorized according to the hierarchical mandatory access principle.

The basis of this approach (the condition for its application) is the introduction of a hierarchy (implemented on any principles) of access objects and subjects and access rights within the compiled hierarchy [11,12.13].

The most typical solution is to introduce a hierarchy of access objects based on the level of confidentiality (public, for official use, no access, etc.) of the documents stored in them, and a hierarchy of access subjects based on their status.

The introduced hierarchy of subjects and access objects allows you to enter permissions (formalize rights) to access resources, for example, after registration, the subject has access with minimal limited rights and privileges (search and view only open collections), the administrator (super administrator) of the upper zero levels has all rights and permissions, the first-level administrator inherits rights and permissions except for some (such as system collections, user data, etc.) and registers other low-level administrators, etc. this principle allows you not to violate the integrity of the system and save information in the system.

The above can be illustrated as follows. Let the sets S = $\{S_1..., S_k\}$ and O = $\{O_1..., O_k\}$ be linearly ordered sets of access subjects (subject) and access objects (object), respectively. As an access subject Si, i = 1..., k is considered as a separate subject, and a group of subjects with the same access rights and privileges, respectively, as an access object O_i, i = 1..., k can also be considered as a separate object and a group of objects characterized by the same access rights to them. Let's introduce the following hierarchy of subjects and access objects: the larger the object's sequence number *i*, the fewer rights and privileges to system resources (documents), respectively, the smaller the subject's sequence number *i*, the more access rights to information it has. Let R = $\{R_1..., Rn\}$ (Rj, where j = 1..., n, (for example, view, edit, link, delete, etc.) be the set of access rights. Where each subject Si uniquely corresponds to Rj a set of rights (privileges) of a set of objects from Oi. For example, subject S1 owns the entire set of rights (privileges) from Rj and has access to all objects of Oi. That is, each subject or group of subject's Si corresponds to a certain set of privileges (privileges) from Rj to certain objects Oi. Where the greater the ordinal number of the subject from *i*, the smaller the set of rights (privileges) from Rj, respectively, has fewer objects from Oi [14,15,16].

Thus, a mandatory access control mechanism is one of the possible ways to implement the principles of rights (powers) to control access to resources. The purpose of the mandate mechanism is to simplify the administration of the system, increase the level of protection of the document during its processing, and allows you to maintain the integrity of the system as a whole.

The basis of the principle of mandatory access control to resources, as noted, is the assignment of access rights of subjects to objects based on their authority. However, in general, several users (subjects) can have the same permissions at the same time, which are taken into account when assigning access rights, for example, several users have the same form of access to work with documents (when setting access control rules (privileges), a set of Rj must be assigned, as a result, these users (subjects) must have the same access rights, i.e. they can work with the same documents). However, the access rights of users (subjects) to documents may differ. Thus, the mandatory access control model implements a fully decentralized principle of organizing and managing the access control process. This approach provides the flexibility to configure the access control system in the database for a specific set of users and resources [17].

III. Technology of Storage Creation and Organization of Access to Information Resources

The developed resource access control model uses a Mandatory Hierarchical Access Control Model. In which each subject has a set of rights (rules) in the system. Each right (rule), in turn, contains a set of privileges that grant access to certain resources. Using elements of the hierarchical access control model. This ensures the decentralization of the management of actors, i.e. the presence in the system of a set of administrators (actors who expose a set of rights (rules) to users) of different levels. A system administrator can delegate some of their authority to lower-level administrators. The system under development supports three main types of actors: actors, a group of actors, and others [18, 19, 20].

The actor – those subjects that are related to the system. Most of the actors in one form or another are users of information resources of the digital library.

A group of actors is a set of actors that have some similar functions when interacting with information resources. And the rest are all registered users who only have access to public collections (search and view information).

In the management of subjects, the following principles are used:

1) Each subject (actor) is assigned classification levels (hierarchy levels), reflecting their place in the corresponding access to objects (collections, documents).

2) Each subject has a set of rules in the system, which is determined by its status in the system being developed. Each of the rights, in turn, contains a set of privileges that grant access to certain resources (objects).

3) Subjects belonging to a certain group inherit a set of rights of this group, i.e. the rights assigned to a group of subjects.

4) Using elements of the hierarchical access control model. The use of this principle means the decentralization of user management, i.e. the presence in the system of many administrators (acting, granting rights to users) of different levels. A system administrator can delegate some of their authority to lower-level administrators. This approach allows, for example, to transfer the responsibility for managing the rights of lower-level entities to a higher level, thereby removing the burden from the system support service.

5) A significant part of the rights of subjects is assigned by their status, or by the subject's belonging to a particular group.

6) The access rights setting is separate from the settings for each of the resources. In other words, the system component that provides the administrator with an interface for editing access rights does not depend on the logic of the resources, including the set of resources itself. Similarly, the resources themselves do not directly depend on the logic of the access control system, which is accessed only by periodic requests to confirm that the current user has a particular privilege.

To describe the operation of the digital library management system, you need to write out the main entities: service (system) collections, collections open for viewing, editing, and closed, metadata contains elements that are divided into the following classes: P-public (for example, DC metadata), A-automatic (owner and date of entry, modifications, etc.), S-service (access rights, class of the metadata element, access to the element, etc.), M-meta (link to pages (dictionary)), U-unique (resource identifier) and each class, in turn, has access to the element (P – all, S-super admin (0 - level), A - admins (1, 2-level).; and documents.

Table I considered the set of privileges and rights of administrative collections divided into five stages.

In addition to the above, the system has system and service collections such as the genre type of the resource (document_type), the data schema (privileges: view and add), user data - privileges (view), on which only the super administrator of the system has rights and privileges [21, 22].

Administrator-level privileges for editing the structure and contents of a collection depend on the attributes of the collection.

If collection access is allowed, then everyone can search and browse the collection, otherwise only the administrator of this collection, the super administrator, and the 1st level administrator.

The administrator can have access to different collections with different privileges to view and edit the collection and edit the records of the collection (documents) as shown in the table.

In addition to the above, the system has system and service collections such as the genre type of the resource (document_type), the data schema (privileges: view and add), user data - privileges (view), on which only the super administrator of the system has rights and privileges [23].

TABLE I. PRIVILEGES AND RIGHTS COLLECTIONS

	Rights to collections							
Privileges	**open**					**closed**		
	1	**2**	**3**	**4**	**5**	**1**	**2**	**3**
1. Collections								
1.1. Create a collection	+	+				+		
1.1.1. Collection data schema (from metadata)	+	+				+		
1.1.2. Document output in the collection	+	+				+		
1.1.3. Displaying (template) a document in a collection	+	+				+		
1.1.4. Sorting a document in a collection	+	+				+		
1.1.5. Structure of documents in the collection	+	+				+		
1.1.6. Description of the collection	+	+				+		
1.1.7. Collection Attributes								
1.1.7.1. Name of the collection	+	+				+		
1.1.7.2. Collection access (open for viewing, open for editing)	+	+				+		
1.2. Delete a collection	+	+				+		
1.3. Access to the collection (allowed)	+	+				+	+	
1.4. Establish a link between collections								
1.4.1. Family relations (according to the content of the collection)	+	+				+	+	
1.4.2. Establishing hard links (links)	+	+				+	+	
1.4.3. Establishing a connection with the rubricator	+	+				+	+	
1.5. Change								
1.5.1. Issuing a document in the collection	+	+	+			+	+	+
1.5.2. Displaying (template) a document in a collection	+	+	+			+	+	+
1.5.3. Sorting a document in a collection	+	+	+			+	+	+
1.5.4. The structure of documents in the collection	+	+	+			+	+	+
1.5.5. Description of the collection	+	+	+			+	+	
1.5.6. Change (attributes)	+	+				+	+	
2. Metadata								
2.1. Create (add) an element	+	+				+		
2.2. Edit an item	+					+		
2.3 Delete an item								
3. Document								
3.1. Create a document	+	+	+	+	+	+	+	
3.2. Edit the document	+	+	+	+	+	+	+	
3.3. Search and view the document	+	+	+	+	+	+	+	
3.4. Delete the document	+	+	+	+	+	+	+	
3.5. Importing a document	+	+	+	+	+	+	+	
3.6. Exporting a document	+	+	+	+	+	+	+	
3.7. Access to the document	+	+				+		
3.8. Change document properties								
3.8.1. Change the document owner	+	+				+		
3.8.2. Change the versioning of the document	+	+	+			+		
3.8.3. Change the document template	+	+				+		
3.8.4. Restrictions on viewing the document (who can see the document)	+	+	+			+		
4. Registration of administrators	+	+	+	+		+		
5. User Privileges								
5.1. Local network access	+	+	+					
5.2. Access to registered users	+	+	+					

Administrator-level privileges for editing the structure and contents of a collection depend on the attributes of the collection.

If collection access is allowed, then everyone can search and browse the collection, otherwise only the administrator of this collection, the super administrator, and the 1st level administrator.

The administrator can have access to different collections with different privileges to view and edit the collection and edit the records of the collection (documents) as shown in the table.

For each administrator or group of administrators in the system, according to the hierarchy level, a certain set of rights and privileges by default is granted to edit the collection and documents (as shown in the table). If necessary, the top-level administrator can add/remove privileges for editing collections and documents. For example, for an administrator who works only with documents, you can give meta description privileges for a collection for an individual collection, or for an administrator who works only with his records in a collection, grant editing privileges to all records of a particular collection [24, 25, 26].

Each administrator-level determines what rights are given to view and edit the collection and what privileges it has, that is:

- permission to view the collection (documents);
- permission to edit the collection, add, edit or delete objects;
- permission to edit collection documents (create, modify and delete a document);
- permission to change the order of collection objects;
- permission to change the format of issuing collection objects;
- permission to add a new object (metadata item);
- Allowing registration of lower-level administrators.

Table II shows the system of hierarchical levels of administration in the system. Administrator-level privileges for editing the structure and contents of a collection depend on the attributes of the collection.

An administrator can have access to different collections with different privileges to view and edit collections and edit collection records (documents). Each collection (depending on the type) has a minimum required set of metadata. Depending on the level of administration, the collection administrator can define a collection metadata schema based on the available metadata from the main catalog [27, 28]. Description of the structure of privileges and rights of administrative collections is shown in Table III.

TABLE II. THE SYSTEM IMPLEMENTS A HIERARCHICAL SIX-LEVEL SYSTEM (MANAGEMENT) OF ADMINISTRATION

Admin Levels	**Has Access**	**Works (Edits)**
0	All collections, documents, schemas, dictionaries, etc.	All collections systems and documents
1	All collections and documents except official ones	All collections and documents except (service and system collections)
2	Only your own and open collections	Structure (meta description of the collection), documents (templates, properties) of the collection record.
3	Open collections	Open collections (meta description of a collection), document structure, and records
4	Open collections	Only records of open collections (also, in some cases, it can work with the structures of individual collections and documents)
5	Open collections	Documents in open collections (only with their records)

TABLE III. RIGHTS AND PRIVILEGES FOR VIEWING EDITING THE STRUCTURE AND DESCRIPTION OF THE COLLECTION

		Rights	Privileges
Collections		Open for viewing,	Search and view a collection
Data schema	Collection metadata	Open for editing	Add an item Edit an item
Meta Description of the collection	Collection Parameters,	open for editing	Change the collection structure, Change the order, Change the format of the document output, Add a new element.
		Open for viewing	Viewing the collection structure
Metadata	Metadata Schema	Open for viewing,	Search and view an item
		Open for editing	Add, Delete, Edit Search and view an item
Document	Documents in the collection	Open for editing	To create, Edit, Remove, Import Export Search and view
		Open for viewing	Search and view a document
	Document Properties	Open for editing	Change the document owner, Change the document versioning, Change the document template
Administrator Registration		Allowed	Add an administrator Remove Administrator

The access rights (privileges) model also allows you to perform the following functions:

- create system administrator accounts;
- creating groups of system administrators with the same default privileges;
- assigning access rights to collections and documents, both for groups of administrators and individual system administrators;
- creating templates for access rights to documents of various types, both for groups and individual users;
- assign access rights to the collection and document structures for the system administrator.

Administrators can be grouped into groups of administrators, which are the following structures:

- name of the group;
- list of administrators;
- a set of rights (privileges) for specific collections and documents;
- a set of rights for document types [29, 30].

In Fig. 3, we discussed the scheme of access to administrative collections and its tasks. Administrator access depends on the level (hierarchy of administrators) of the administration and the attributes of the collection. As noted above, in the system, collections can be divided into Administrative and Content Collections. Each Content collection consists of documents. The document, in turn, contains Information Content, an Annotation, and a Link (that is, a pointer to the data warehouse (repository) where the full text is stored, to an archive file, a separate fragment (figure, table, etc.). The document also has a Meta Description, Genre Document Type, Versioning, and Record Owner. The meta description of the collection documents, in turn, has a Template, Structure, Rules for ordering and displaying the document.

Document versioning is a document that corresponds to a certain stage (stage) of document development. Each version represents the state of the document at some point in time. Users can only view the full versions of the document. Document management is the process of tracking such information, ensuring, for example, that there is only one original copy of a document and that archived copies of all previous versions are stored in the order in which they appear [31, 32].

Abstract - A brief description of the document, explaining its content, purpose, form, and other features. Annotation, information about an object for which the main purpose is to annotate the "target" Resource (or its scope). Examples of such Annotation Objects include notes, structured comments, and links. Annotations help in interpreting the target resource, or about support or objections, i.e. more detailed explanations. A relationship is a relationship between an entity and what is associated with it.

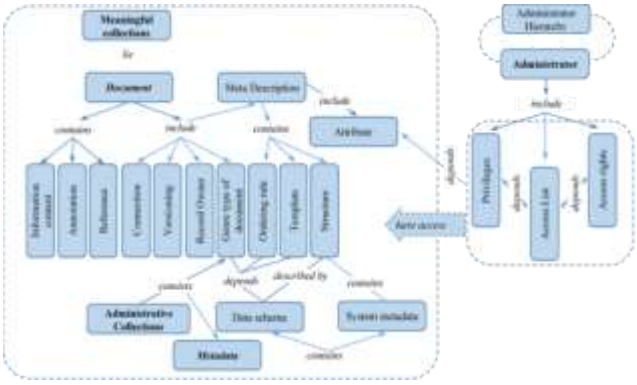

Fig. 3. System for Managing Administrator Access to Collections.

Connections — This is what the Digital Library consists of. The relationships between the data are identified by the type of relationship and the direction. The nature of the type and direction of the relationship can be reflected by its name (named relationship, relationship characteristics). Relationships allow objects in one collection to link to other objects not only in a given system but also to other remote distributed systems. There are two types of communication (relationships):

- Internal relationships – An order relationship between documents that builds a hierarchy of subordination in a collection, for example, a relationship of subordination between documents;
- External relations – The relationship between documents regardless of the structure of subordination.

A reference is an entry in a document that points to another part of that document or another document. A link can point to a data warehouse (repository) where the full text is stored, to an archive file, or a separate fragment (figure, table).

Ordering rule – The system also arranges the list of documents according to certain weights that are assigned to the documents during the search. Cataloging implements the basic paradigm of organizing information and ensures its search by pre-defined criteria (for example, by list size, alphabetically) [33, 34].

The owner of the record is the actor who creates, modifies documents in collections, and establishes relationships between them. The owner of the document record is responsible for the storage, correctness, and presentation of the document by the users of the system. Also, the document may not belong to the system, i.e. its “owner” may be another information system, and our system contains only its description or a link to this document.

Attribute - A feature that characterizes an object or entity, its properties, a data descriptor that contains one of its characteristics (for example, name, type, access elements, presentation form, etc.).

Structure - A fixed ordered set of objects and relationships between them (information about its logical division).

A template is a model for creating a collection document. The template stores a variety of elements that make up the basis of the document: the structure of the document together with the attributes assigned to them; document page parameters; a list of available styles; macros (a sequence of actions that automate working with the document); custom toolbars. Document template describes the basic structure of the document, which is processed and filled with real data by the filling function. To create a new collection, you must primarily use document templates to ensure that the documents are designed under the Organization's accepted style [35, 36].

Fig. 4 shows a picture of changing the characteristics of administrative colleges.

Each available collection of the administrator contains a meta description as shown in Fig. 4 above (edit) and delete where it has the right to search for documents, view documents, enter a new document, and a meta description of the collection. In the meta description of a collection, you can change the order of objects (the document display rule), change the collection parameters (i.e., the attributes are closed or open for viewing and editing), change the format of object output (the collection structure), add a new object (add a metadata element), and edit the information in the collection (edit the document).

Each available collections of the administrator contains a meta description, as shown in Fig. 5 above (edit) and delete where it has the right to search for documents, view documents, enter a new document and a meta description of the collection.

Editing a document has a minimum set of privileges such as shown in Fig. 6: add (entering a new document, importing a new document), view (show), edit/edit (edit), establish a link (link) and delete a document (delete). Depending on the access of the collection and the structure of the collection itself, these privileges may be minimal.

Collections: Persons in the collection of computer science training courses (persona_cat), Mixed learning: Key terms (term_cat), List of sections of the course "Modern problems of Computer Science" (temas_class) minimum set of privileges.

Fig. 4. For Example, Admin Page.

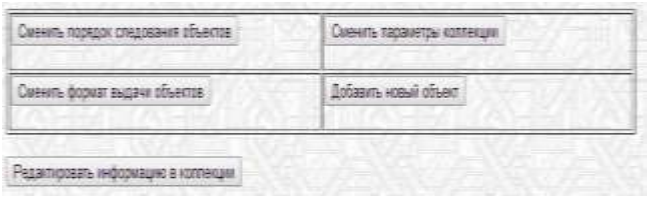

Fig. 5. Meta Description of the Collection.

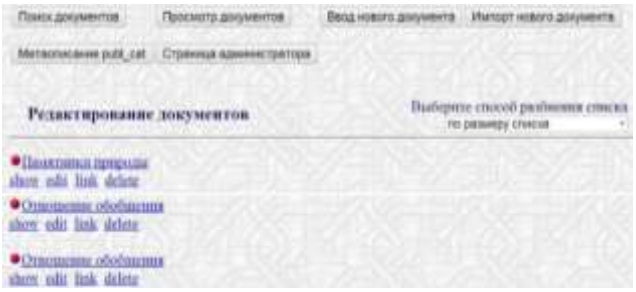

Fig. 6. For Example, Resources in the Computer Science Course Collection (publ_cat) contains the above Privileges.

The system implements a hierarchical six-level system (management) of administration, as shown in Fig. 7.

Each of the administrators also has the right to view open collections, register other lower-level administrators, and have all the rights and privileges of low-level administrators.

This access control method allows you to control user access to information based on the status in the system. The use of this method involves determining the rights (rules) in the system. The concept of law (rules) can be defined as a set of actions and powers related to a certain type of activity. So, instead of specifying all the access types for each user to each object, it is enough to specify a set of object access rules. And users, in turn, specify their privileges [37].

Rights are a set of rules (rights) that determine what privileges and over what objects the user who is assigned this set will have. Rules that make up a set of rights can be permissive (allow acting on a certain group of objects) or forbidding (prohibit performing the corresponding actions).

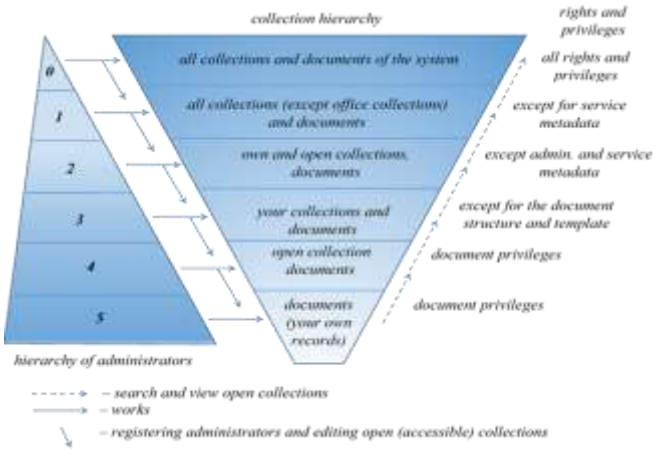

Fig. 7. Collection Hierarchy.

A privilege is all possible actions of a user (administrator) to perform any actions about objects (collections and documents) of the system, which requires a certain differentiation of access to this action. In principle, it is possible to combine several such actions into one privilege, but only if this combination of the simplest actions is required in all possible situations.

In other words, the user can perform different actions in different situations. The same set of rights (rules) can be used by several different users. In the system, a user can also have multiple rights (rules) at the same time.

The main advantages of this access control model are:

Easy administration – separation of rights (rules), privileges, and users allows you to split the task into parts: defining user rights and defining access rights to the object to perform certain functions. This approach simplifies the administration process since when changing the user's area of responsibility for objects, it is enough to change his previous rights (rules) and assign other appropriate ones for this situation [38].

Hierarchy of rights - the system of rights (rules) can be configured so that it reflects the real processes much more closely by building a hierarchy of rights. Each right (rule), along with its privileges, can inherit the privileges of other rules. This approach also significantly simplifies the administration of the system.

The principle of least privilege allows the user to register in the system with the least rights, allowing him to perform limited tasks. Users who have multiple rights do not always need all of their privileges to perform a specific task.

The principle of least privilege is very important to ensure the integrity of the data in the collection. It requires users to be given only those of the privileges allowed to it that it needs to perform a specific task.

To do this, you need to find out the goals of the task, the set of privileges required for its execution and restrict the user's privileges to this set. Prohibiting user privileges that are not required to perform a certain task allows you to maintain the integrity of the data in the system [39].

Separation of responsibilities is one important principle in the access control system. The system of mandatory access control helps to solve this problem with maximum simplicity. The system of mandatory access allocation allows you to distribute the problems of sharing responsibility between administrators and privileged users of the system and provides a multi-stage security level of the entire system.

The model consists of the following entities: users, sets of rights (rules), and privileges. A user is either a person or a program running on behalf of the user. This includes not only people but also external (relative to the digital library) software and hardware.

Thus, the advantage of this model is the ease of administration: assigning users a set of access rights (rules). At the same time, it does not allow you to manage different parts of the system separately and even more so - to delegate such powers to any user - ensures the integrity of the system. This model combines the protection and restriction of rights applied to data, collections, and system collections designed to prevent their unwanted use.

Classification of the hierarchy of entities and objects (resources) have rights and permissions (such as creating, editing, deleting, registering, controlling access, etc.)

After identification and authentication, each Sk user (subject) is assigned a minimum set of rights (privileges) Rj to objects from the Oi (for example, search and view documents that are open for viewing). Further, depending on the work in the system or the status of the administrator (subject), the top-level administrator grants the rights (privileges) to the corresponding level of administration to the corresponding objects, etc. or vice versa. This control method, as noted above, allows you to control the operation of the system and preserves the integrity of the system as a whole. Also, top-level administrators give part of their authority to lower-level administrators and thus freeing themselves from a large load of work in the system [40].

IV. Conclusion

The article is devoted to the history of the creation and description of the technological solutions used in the creation of the system. The article describes the architecture of the information system and the principles of integration with external sources, the rules for the representation and transformation of metadata, as well as the work with dictionaries that are used to systematize and classify information resources, and modeling the relationships between them.

An important component of the EC is a digital repository or repository of digital objects, the main purpose of which is to store these objects with all their possible variants and versions. In a broader sense, a digital repository is not only a repository of digital objects but also a system that provides a service for both managing these objects and organizing access to these objects. The International Organization for Standardization (ISO) has proposed the ISO-14721 (Open Archive Information System – OAIS) standard for organizing a system of long-term storage of information resources (a repository of digital objects) [8]. The reference model for the OAIS standard is a conceptual model based on the extended DublinCore data schema (Dublin Core, Dublin Core, DC), and is a set of metadata elements for representing the domain ontology.

Acknowledgment

This research has been funded by the Science Committee of the Ministry of Education and Science of the Republic of Kazakhstan (Grant No. AP08857179).

References

[1] Shokin Yu. I., Fedotov A.M., Guskov A. E., Zhizhimov O. L., Stolyarov S. V. Digital libraries - the way of integration of information resources of the Siberian Branch of the Russian Academy of Sciences. Series: Mathematics, Mechanics, Computer science. - 2005. - Almaty: KazNU. - No. 2. - S. 115-127. - ISSN 1563-0285.

[2] Law of the Russian Federation (No. 149-FZ) "On information, information technologies and information protection" dated July 27,

2006 // - the Federal law dated 27 07 2006 N 149-FZ (as amended on 07 06 2013). - 2006.

[3] GOST ISO / IEC 2382-1-99. Information technology. Dictionary. Part 1. Basic terms and definitions / / - Minsk: Mezhgos. council for Standardization, Metrology and Certification. - 1999. - 40 c.

[4] GOST 34.320-96. Information technologies. Database standards system. Concepts and terminology for the conceptual scheme and information base / / - Gosstandart of Russia, M.: IPK Publishing House of Standards. - 2001. - 14 p.

[5] GOST 7.0-99. The system of standards for information, library and publishing. Information and library activities. Bibliography. Terms and definitions.

[6] Golitsyna O. L., Maksimov N. V. Information systems / Moscow International Institute of Econometrics, Informatics, Finance and Law. - Moscow: 2004. - 329 p.

[7] Izbachkov Yu. S., Petrov V. N. Information systems: Textbook for universities. 2nd ed. - St. Petersburg: Piter, 2006. - 656s.

[8] Karminsky A.M., Chernikov B. V. K24 Information systems in the economy: In 2 ch. Ch. 1. Methodology of creation: Textbook. - M.: Finance and Statistics, 2006. - 336 p.: ill.

[9] Soviets, B. Ya. Information technologies: textbook for universities / B. Ya. Soviets, V. V. Tsekhanovsky. - M.: Higher School, 2005.

[10] Kogalovsky M. R. Perspective technologies of information systems. - M.: DMK Press; IT Company, 2003. - 288 p.

[11] Berg A. I. "Questions of cybernetics", VK-72 / Ed. by R. M. Suslov and A. P. Reutov. - M.: Scientific Council of the USSR Academy of Sciences "Cybernetics".

[12] Big systems and management (ed. by V. I. Chernetsky). Izd. LVVIK im. A. F. Mozhaisky, Leningrad, 1969. 206 p.

[13] Telemtaev M. M. Information systems. Moscow: MTS, 2010, 98 p.

[14] Kutsenogiy K. P., Kutsenogiy P. K., Molorodov Yu. I., Fedotov A.M. Development of the metadata structure for atmospheric aerosols based on an information model. 2004. Vol. 9. Special Issue: Proceedings of the International Conference "Computing and Information Technologies for Environmental Sciences" (CITES 2003). Tomsk, September 1-11, 2003 Part 2. pp. 25-33.

[15] Shokin Yu. I., Fedotov A.M., Zhizhimov O. L., Fedotova O. A. Management system of digital libraries [Digital resource] / / Distributed information and computing resources (DICR-2014): materials of the XV Russian Conference with international participation (digital edition). - 2014. - Novosibirsk: Institute of Computing Technologies of the Siberian Branch of the Russian Academy of Sciences. - State Register. No.: 0321500379.

[16] Larkov N. S. Documentovedenie: Uchebnoe posobie / N. S. Larkov. - M.: AST: Vostok-Zapad, 2006.

[17] Pavlenko N. A. Istoriya pisma. Minsk, 1987.

[18] Larkov N. S. Documentovedenie: Digital textbook. Tomsk: TSU 2002.

[19] Otle P. Biblioteka, bibliografiya, dokumentatsiya: Izbrannye trudy pionera informatiki [Library, bibliography, documentation: Selected works of the pioneer of Informatics]. - Moscow: FAIR-PRESS: Pashkov House, 2004. - 348, [1] p.- (Special publishing project for libraries). - Bibliogr.: pp. 312-327. - Names. decree: pp. 340-342. - ISBN 5-8183-0624-0.

[20] GOST R ISO / IEC TO 10000-2-99. Information technology. Fundamentals and taxonomy of functional standards. Part 2. Principles and taxonomy of VOS profiles.

[21] Shvetsova-Vodka G. N. General theory of the document and the book: textbook. manual / G. N. Shvetsova-Vodka. - M.: Rybari ; K.: Knowledge, 2009. - 487 p.

[22] Stolyarov Yu. N. "There is no alternative to the document" / Yu. N. Stolyarov / / Nauch. and tech. b-ki. – 2000. – № 3.

[23] Stolyarov Yu. N. The theory of relativity of the document / Yu. N. Stolyarov / / Ibid. – 2006. – № 7.

[24] Sokolov A.V. Social communications: textbook. - method. manual: in 2 ch. Ch. I. / A.V. Sokolov-Moscow: Profizdat, 2001. - 222 p. - ("Modern library". Issue 16).

[25] Dvoinosova G. A. / Functions of the document / / Scientific and technical information. Ser. 1: Organization and methodology of information work. - 2013. - № 2.

[26] Kushnarenko N. N. Documentovedenie: ucheb. / N. N. Kushnarenko. - 2nd ed., pererab. and additional – Kiev: Knowledge, 2000. - 459 p.

[27] V. V. Pogulyaev Commentary to the Federal law "On the mandatory copy of documents".

[28] Tkachev A.V. The legal status of computer documents: main characteristics. M., 2000.

[29] Ozhegov S. I., Shvedova N. Yu. Explanatory dictionary of the Russian language. M., 1993. P. 241.

[30] Oblavets A. A. Legal regulation of digital document management//Vestnik MSU. Ser. 11. Pravo. 1997. No. 4, p 51.

[31] Walk, V. V., E. A. Morgunova Comments to the Federal law "On information, Informatization and protection of information". M., 2004, P. 14.

[32] Federal law of December 29, 1994 N 78-FZ "On librarianship" (with amendments and additions).

[33] Federal law No. 217354-4 "On information, information technologies information security".

[34] Federal law of October 22, 2004 No. 125-FZ "On archival business in the Russian Federation" (with amendments and additions).

[35] GOST R 52292-2004 Information technology. Digital exchange of information. Terms and definitions.

[36] GOST R 52292-2004 Information technology. Digital information exchange Terms and definitions.

[37] Fedotov A.M. Methodologies for constructing distributed systems / / Selected reports of the X Russian Conference "Distributed Information and Computing Resources" (DICR-2005), Novosibirsk, October 6-8, 2005 / Computational Technologies. 2006. - Vol. 11. - p. 3-Novosibirsk: IVT SB RAS. - ISSN 1560-7534.

[38] Kogalovsky M. R. Metadata in computer systems/M. R. Kogalovsky / / Programming, 2013, N No. 4. – p.

[39] 28-46 39. Kogalovsky M. R. Scientific collections of information resources in digital libraries. Proceedings of the First All-Russian Scientific Conference "Digital Libraries: Promising Methods and Technologies, Digital Collections", St. Petersburg, October 1999. St. Petersburg University Press, 1999.

[40] Kogalovsky M. R. Metadata, their properties, functions, classification and means of representation // Proceedings of the 14th All-Russian Scientific Conference "Digital Libraries: Promising Methods and Technologies, Digital Collections" - RCDL-2012, Pereslavl-Zalessky, Russia, October 15-18, 201.