РАЗРАБОТКА ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ И ИНЖЕНЕРИЯ ЗНАНИЙ

УДК 621.38; 004.9
МРНТИ 28.23.33; 28.17.31; 44.37.29; 44.37.03

INFORMATION SYSTEM OF DATA STORAGE CENTER

G.I. KHASSENOVA, D. NURADIL, D. KHAIROLLA, A. TOLEGENOVA

International University of Information Technologies

***Abstract**: The problem of increasing and processing the volume of archival documents today is relevant for many organizations and institutions. Archives accumulate for years, tens of years, for which huge volumes are collected. This leads to significant difficulties in the organization of modern archives. A modern and effective solution to this problem is the creation of an electronic archive (EA). Moreover, the development of modern technology suggests a gradual, increasing importance of electronic archival systems. The system can help to spend less resources and upgrade data storage efficiency. Compared to traditional storage systems with simple algorithms, archival data storage has more advantages. In this article we have presented a new approach to the storage and processing of archival data with minimal expenditure of resources of the Institute.*

***Keywords**: Archive. Data base. Information system. Web platform. Structuring. The center of the store. Management of the archive. Stack. Cells*

ҚҰЖАТТАРДЫ ОРТАЛЫҚ САҚТАУДЫҢ АҚПАРАТТЫҚ ЖҮЙЕСІ

***Аңдатпа:** Мұрағат құжаттарының өсу және көлемін өңдеу мәселесі бүгінде көптеген ұйымдар мен мекемелер үшін өзекті болып отыр. Мұрағаттар жылдар, ондаған жылдар бойы жинақталып, үлкен көлемдер жиналады. Осыған орай қазіргі заманғы мұрағат қоймаларын ұйымдастыру кезіндегі елеулі қиындықтарға әкеледі. Бұл проблеманың заманауи және тиімді шешімі электрондық мұрағат (ЭА) құру болып табылады. Сонымен қатар, қазіргі заманғы технологияларды дамыту электрондық мұрағат жүйелерінің біртіндеп, өсіп келе жатқан мәні туралы айтуға мүмкіндік береді. Жүйе аз ресурстарды жұмсауға және деректерді сақтау тиімділігін жаңғыртуға көмектесе алады. Мұрағаттық деректерді сақтау жүйесі, қарапайым алгоритмдермен дәстүрлі сақтау жүйелерімен салыстырғанда артықшылық алады. Бұл мақалада біз институт ресурстарын аз жұмсай отырып, мұрағат деректерін сақтау мен өңдеудің жаңа тәсілін ұсындық.*

***Түйінді сөздер:** мұрағат, деректер қоры, ақпараттық жүйе, веб-платформа, құрылымдау, сақтау орталығы, мұрағат басқармасы, сөре, ұяшықтар*

ИНФОРМАЦИОННАЯ СИСТЕМА ЦЕНТРА ХРАНЕНИЯ ДАННЫХ

***Аннотация:** Проблема возрастания и обработки объёмов архивных документов сегодня актуальна для множества организаций и учреждений. Архивы накапливаются годами, десятками лет, за которые собираются колоссальные объемы. Это приводит к значительным трудностям при организации современных архивохранилищ. Современным и эффективным решением этой проблемы является создание электронного архива (ЭА). Более того, развитие современных технологий позволяет говорить о постепенном, возрастающем значении электронных архивных систем. Система может помочь потратить меньше ресурсов и модернизировать эффективность хранения данных. По сравнению с традиционными системами хранения с простыми алгоритмами, система хранения архивных данных*

имеет больше преимуществ. В данной статье мы представили новый подход к хранению и обработке архивных данных с минимальным расходованием ресурсов института.

Ключевые слова: *архив, база данных, информационная система, веб-платформа, структурирование, центр хранилища, управление архива, стеллаж, ячейки*

iNTRODUCTION

Nowadays, modern information technologies have reached a high level of development. Big data plays a big role in our lives. Every year there are more and more documents in the world. And the resources for their storage are increasing, and the search for documents over the past years is also becoming more complex with the arrival of new documents. Note that at the moment technologies allow storing large amounts of data in electronic format. But, unfortunately, many institutions use the old method of storing document data in paper format.

The practical base of the research is presented by the Institute of archaeology named after A. H. Margulan. The Institute provides an opportunity to develop an internal system or public Internet service for long-term storage of documents with a high guarantee of document safety, and minimal risk of data loss due to unexpected failure of hardware or software.

Figure 1. Place of research

Currently, the Institute of archaeology does not have an electronic system for adjusting the archive. The current problem of this organization is that they store documents on paper or hard copy (non-digital) by organizing their own archives or archives of local importance. Storage of documents in hard copy complicates the process and increases the search time of documents if necessary to reproduce it, complicates the recycling process, increases the area allocated for the archive, occupying more and more square meters in the room, when this area can be used for more useful purposes. The process of storing documents in a hard copy is not an easy task, as there are certain rules of storage, given the room temperature, humidity, fire safety, size and material of windows and doors, the distance between the shelves, etc. the Search for documents is manual, i.e. the employee of the Institute of archaeology is looking for a certain stack first, he is looking for the necessary catalogue, and in this catalogue finds the document he needs.

Storage of documents in this Institute is carried out in an orderly format. The document storage archive is as follows. Archive documents are stored in stacks (8), there are cells in each stack. Each cell contains boxes or albums. Documents are divided into two types – documents of the Soviet period and documents of the modern period. The storage scheme is shown in figure 2.

To create an inventory number for the object stack numbers are used from №901,902...908(for each). For numbering cells to these inventory numbers added letters of the Latin alphabet, for example, №901-A. Numbering begins on the left side of the storage and from the top shelf of the stack. In boxes are stored documents of the Soviet period, in albums documents of the modern period. Each box has its own inventory number. The first digit indicates the number of cells in the archive 8 stacks and each has 8 cells - 8×8=64. The inventory number of boxes starts with №1,2..64;

STRUCTURE OF INFORMATION SYSTEM

The box stores about 300-350 data. There are four boxes in one cell, in the total amount of 1000-2000 documents-data. These are field reports, field diaries, field albums with photos, extracts, historical references, drawings. The inventory number of the box looks like this:

№1.100-226-901-А. (1)

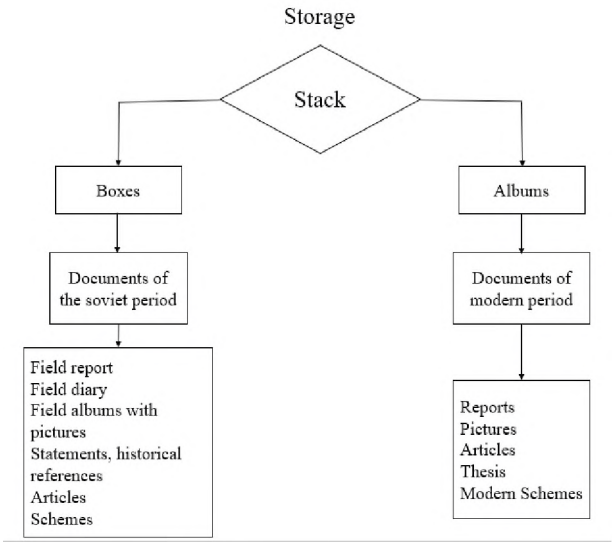


Figure 2. The structure of storage of documents at the Institute of archaeology named after A. H. Margulan.

As indicated on the inventory number under number 1, 1 - box number, 100-226 numbering of documents, 901-A stack and cell. Documents of the Soviet period are classified by year. This Institute has existed since 1949. The earliest documents have a category of 100-500.

We have created a repository of archival data – a center that provides structured storage of archival documents. It also includes document management, provides migration of archival documents. The purpose of the system is to automate the basic data processes of the Institute of archaeology in accordance with the regulatory and methodological documents in the field of archives.

The information storage system will allow storing, deleting and supplementing archival documents to the staff of the Institute of archaeology. An employee of the organization, if necessary, finds the necessary document, enters the data storage system, enters the name of the archive document in the search, and as soon as the employee finds the desired file, it can be opened, downloaded to a computer in PDF format.

A staff member of the Institute finds the document he needs through a search engine on the website. To search for information using a search engine, the user formulates a search query. The job of a search engine is to ask the user to find documents containing keywords or words as any related keywords. The search engine generates a search results page. These search results can contain different types of results, such as images, audio files, web pages, and documents.

The main components of the search engine: search robot, indexer, search engine. As a rule, the systems work in stages. First, the search engine receives the content, then the indexer generates the index to be indexed for search, and finally, the search engine provides the functionality to search for indexed data. To update the search engine, this indexing cycle is repeated.

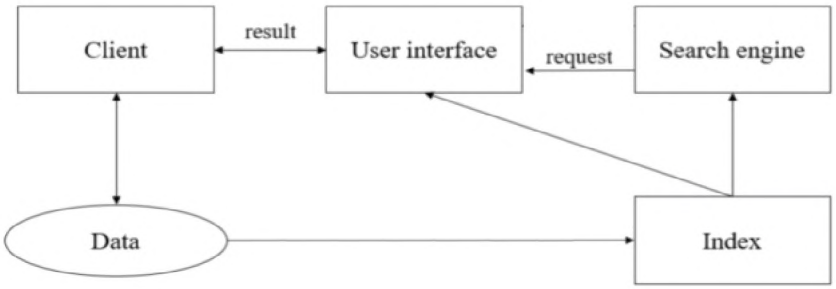


Figure 3. The process of search engine

RESULTS

The archival data storage system is currently used in many States. Since it is necessary to work with a considerable amount of documents, their translation into electronic form greatly facilitates the process. Actively developing companies are faced with a considerable turnover of documents. The essence of the formation of databases is to prepare electronic copies of all available paper documents by scanning or automated input.

As a rule, a local server is used as a data storage for all archival documents of the Institute. PHP allows you to use a variety of database management system, but the most popular today in conjunction with PHP is MySQL. MySQL provides free software that allows you to interact with databases using SQL commands.

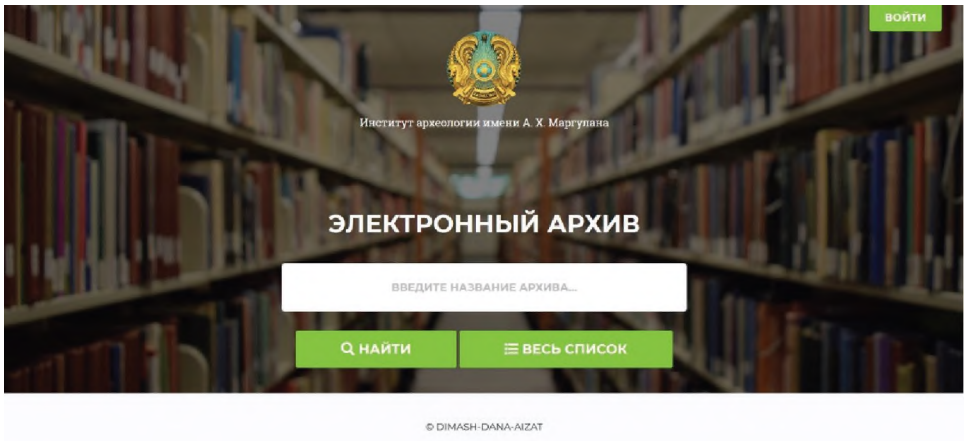


Figure 4. Main page

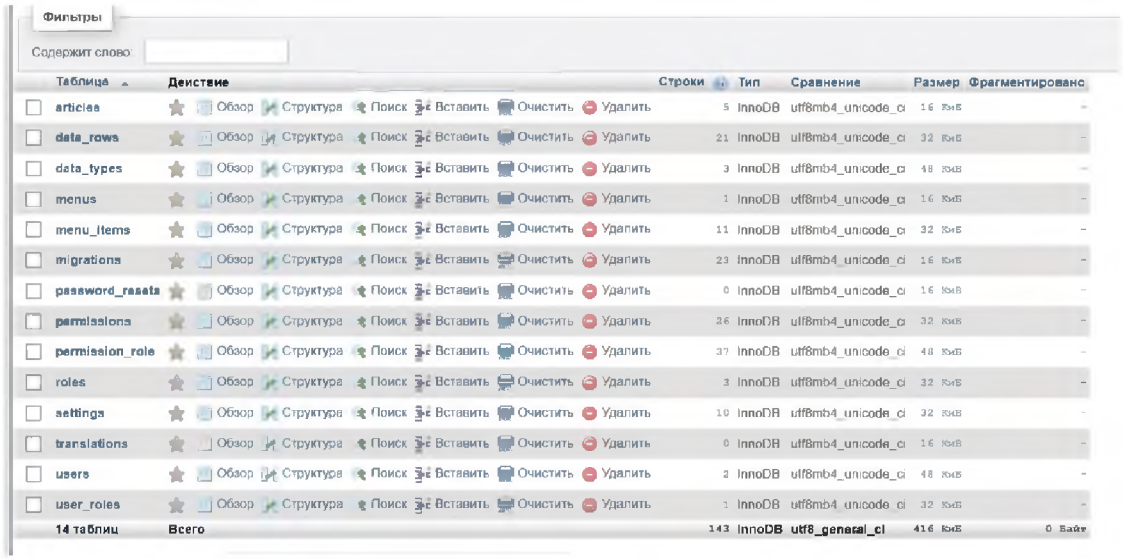


Figure 5. Website database management via phpMyAdmin web application

To simplify working with MySQL databases, a special set of phpMyAdmin scripts is installed. phpMyAdmin presents an intuitive web interface for managing MySQL databases. Using this tool, it is much easier to work with databases than to manage MySQL through the console.

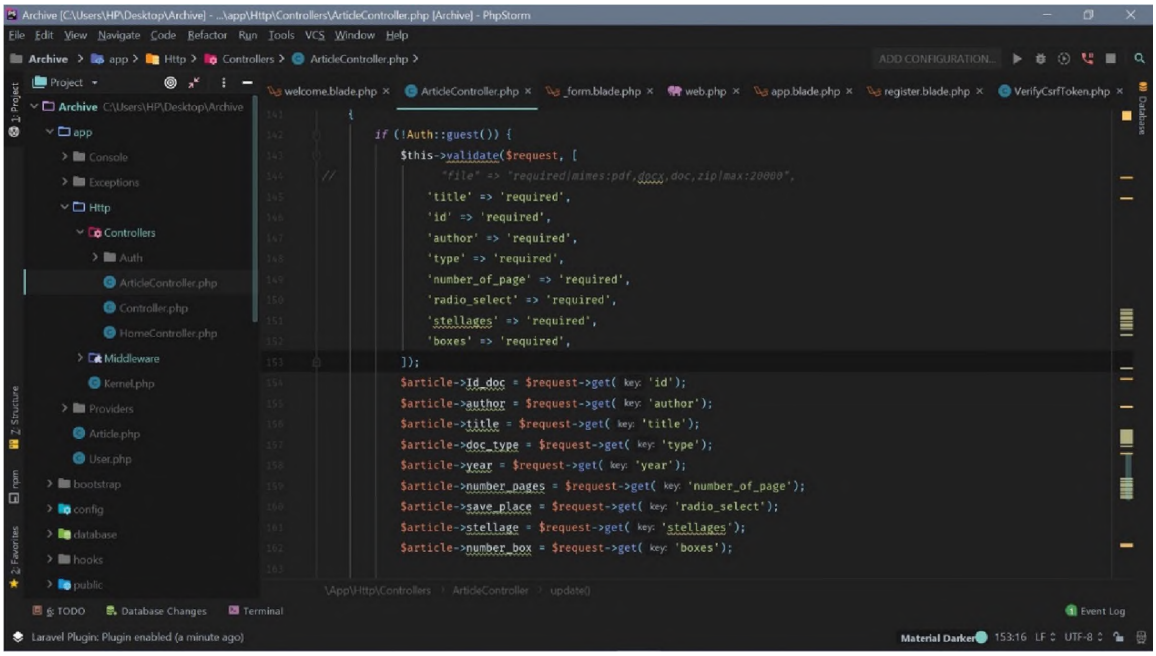
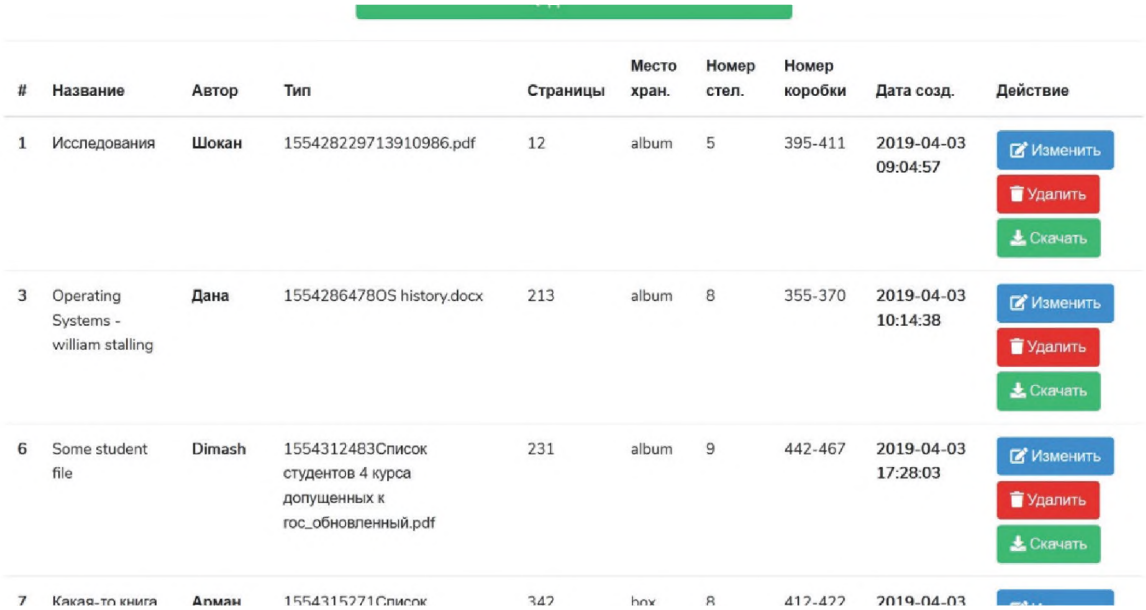


Figure 6. Part of the program code of the server part of the website

The search for documents goes by hand, i.e. an employee of the Institute of archaeology first looks for a certain rack, it looks for the necessary directory, and in this directory finds the document he needs.



#	Название	Автор	Тип	Страницы	Место хран.	Номер стел.	Номер коробки	Дата созд.	Действие
1	Исследования	Шокан	155428229713910986.pdf	12	album	5	395-411	2019-04-03 09:04:57	Изменить Удалить Скачать
3	Operating Systems - william stalling	Дана	1554286478OS history.docx	213	album	8	355-370	2019-04-03 10:14:38	Изменить Удалить Скачать
6	Some student file	Dimash	1554312483Список студентов 4 курса допущенных к гос_обновленный.pdf	231	album	9	442-467	2019-04-03 17:28:03	Изменить Удалить Скачать
7	Какая-то книга	Арман	1554315271Список	342	box	8	412-422	2019-04-03	

Figure 7. Displaying data in graphical form on the website

The effectiveness of the information system depends largely on its architecture. Currently, the client-server architecture is promising. In a fairly common version, it involves the presence of a computer network and a distributed database, including a corporate database and personal database. The corporate database is located on the server computer. The personal database is located on the computers of employees of departments who are clients of the corporate database.

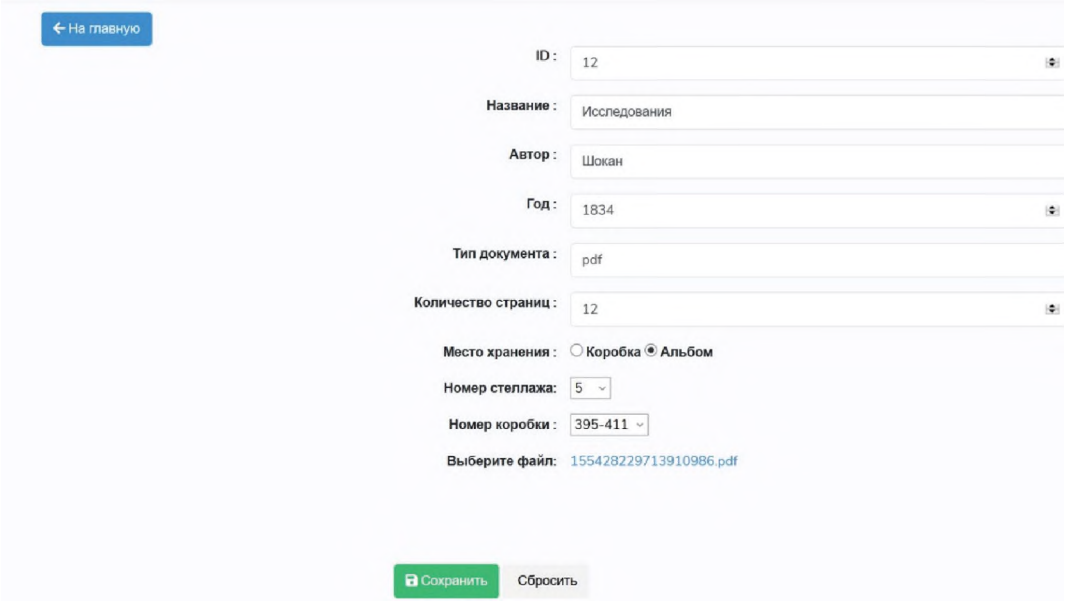


Figure 8. Fields to add information to the new archive

CONCLUSION

The concept of «archive» today does not mean dusty shelves filled with documentation, and electronic systems that can store data for a long time. Despite all the advantages of specialized professional systems, the way of storing this or that information, especially in Kazakhstan, does not always meet the real requirements. This article introduced a sophisticated system for storing archived data. The proposed approach simultaneously performs several solutions for the management of archival documents of the Institute named after A. Margulan. The first and most important requirement for an electronic archive is the exclusion of the physical possibility to delete or modify data both by negligence and malice. In other words, the information carrier must provide a single record when reading multiple times. As a result, data protection against deletion should be software-based. In addition, storage durability is a key requirement.

REFERENCES

1. G.A. Egorov, V.I. Šiaudkulis, M. Finotti, M.I. Belyakov. Principles of practical implementation of modern archive data warehouses, 2014, no.11, pp. 54-57. (in English)
2. A.A Sorokin, S.P. Korolev, S.I. Smagin, A.N. Polyakov. Layout of a fault-tolerant information system for cloud storage of scientific data sets,2012,no.8,pp.192-193(in Russian)
3. Greg Schultz. Cloud and Virtual Data Storage Networking, 2011, no.3, pp.30- 35. (in English)
4. John William Toigo. The Holy Grail of Data Storage Management,2009,n.16,99.154-158. (in English)
5. Raymond Camden. Client-Side Data Storage: Keeping It Local,2013,no.9,pp.284-287. (in English)
6. Г.И. Хасенова, А. Шойынбек. Выбор облачной платформы на основе требовании к центральному хранилищу электронных документов. Международный научный журнал «Весник КБТУ» №3(39) г. Алматы 2016.
7. Г.И. Хасенова, Т. Темирболатова, У. Темирболатова. Технология интеграции больших неоднородных данных. Труды II международной научно-практической конференции «Информационные и телекоммуникационные технологии: образование, наука, практика». 2015. КазНИТУ имени К.Сатпаева.