



# THE PROCESSING'S AUTOMATION OF DIGITAL DOCUMENTS FOR HYPERTEXT SCIENTIFIC LIBRARY

Alexey Alexandrovich Nedelkin, Valery Alexandrovich Titov, Elena Ivanovna Tikhomirova and  
Yuliya Dmitrievna Romanova  
Plekhanov Russian University of Economics, Stremjannyj str., Moscow, Russian Federation  
E-Mail: [aa@nedelk.in](mailto:aa@nedelk.in)

## ABSTRACT

The problem of processing of considerable number of digital images of the electronic copies of scientific works which contain the text in the automatic mode for formation of scientific works' electronic copies and formation of digital scientific library on the basis of the site of university is investigated in the work. The experience of creation and the use of hypertext electronic library resources is analysed and generalized in the work. The basic concepts which differentiate the ideas of the traditional and electronic document for more exact characteristic of hypertext technologies in education are given. The ways and methods of the solution of complex challenges of the automated transformation of the scanned images to an electronic format are shown. The approaches for the improvement of the quality of the scanned materials are described.

**Keywords:** hypertext, hypertext materials, electronic educational editions, library and information educational environment, transference of the documents in an electronic form.

## 1. INTRODUCTION

Modern universities constantly improve and introduce new technologies and methods to training in the educational process. One of the technologies is a hypertext. The hypertext is an attractive format for the publication because of the flexibility of the final product, and also it gives different opportunities for publishers, researchers, scientists and students because the cost of the replication of such production aspires to a minimum [1].

Initially the technology of the hypertext was used to systematize various information and documents, to improve the use of big information resources, and also to coordinate complex challenges of planning. Well-structured documents in a format of the hypertext give the chance of the fast search and instant access to the necessary information and data [2].

The main idea of hypertext technologies is that that the search of documentary information is realized taking into account a set of the interrelations which are available between the documents, so, more effectively than at traditional methods of search.

The access to full texts of scientific works and other materials can be carried out both while using a computer, and using mobile devices on the Internet according to the passwords. It gives the chance to unite the readers who are from different parts of the world in one information space and to limit the access for the third parties at the same time [3].

In comparison with traditional libraries hypertext electronic and digital libraries have a number of advantages, allowing the libraries and readers to solve a number of important problems: the creation of specialized collections for the ensuring of educational process or research; reduction of material damage from the careless use or loss of printing materials; to adapt for the needs of the reader; to inform the readers about the last receipts in a library stock [4].

Modern technologies give an opportunity to make a reading process similar to the work with paper thanks to: the supports of a full-fledged color rendition; a creation of bookmarks and notes in a source directly; browsing of pages with a soundtrack, etc. [5].

## 2. METHODS AND OPERATING PROCEDURE

### 2.1 The research of a problem

The comparative analysis of traditional forms of education with the application of modern multimedia means of teaching in the USA and Europe showed the advantages of the hypertext materials, and also in the creation of selections which are interesting for the users. The existing approaches to the automation of a library science were studied and analysed in the course of the analysis of modern information technologies and requirements of the Russian legislation [6].

### 2.2. The solution

Hence, the decision on creation of our own hypertext electronic and digital scientific library which corresponds to the set of requirements was made. Web - oriented architecture of hypertext electronic and digital scientific library was projected, proceeding from the existing principle of work of a corporate network at the University [7]. This architecture allows to use any modern browser as the client of the hypertext electronic and digital scientific library, thereby, solving a problem of the compatibility and uniformity of the software.

### 2.3. The description of the decision

The choice of the development's technologies was predetermined by the chosen implementers used in the organization that allows: to carry out the fullest integration of the hypertext electronic and digital scientific library with the main portal of the organization, and also it is essential to save the appliances as there is no need to buy



the additional software for the work of hypertext electronic and digital scientific library of the organization [6]. Thus, the developed hypertext electronic and digital scientific library of the organization can be entirely integrated into the main portal of the university and can be used with the existing equipment and the software.

The hypertext electronic and digital scientific library consists of three levels: the representation and management of one or several web - servers; the logic of the electronic library which is operated by an application server; the storage of data, the basis of which is the database management system. The main logic of the REU hypertext electronic and digital scientific library of G. V. Plekhanov is realized in the PHP language and is carried out on the server in the form of the interpreted files with an open source code [6].

The storage of data is on the separate server and represents a relational database; business- logic is on the separate web - server in the form of the interpreted initial code.

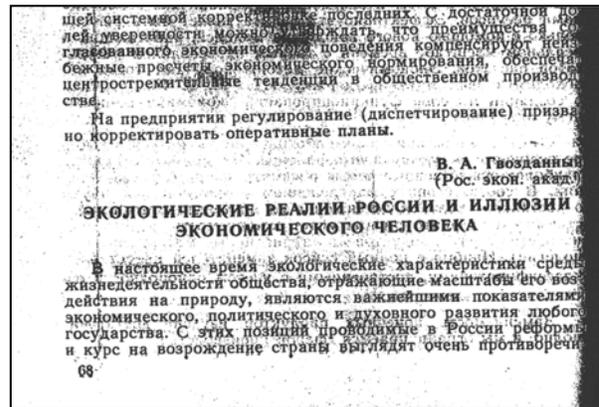
#### 2.4. Tools

The system is cross-platform and at present functions under the control of Windows Server 2012 OS. Microsoft IIS is used as the web - server, and DBMS MySQL 5 is used as DBMS.

#### 2.5. The adaptation of the results of research to real requirements

It is necessary to fill a number of requisites for the publication on the site of the electronic copy of the paper edition: authors, the name of the work, the place and the year of the edition, etc. During the creation of the scanned books the portal uses files in the PDF format as a source of publications [8]. Further a number of specially created scenarios and algorithms are started on the server in a completely automated mode.

A manual editing of images by the means of habitual graphic editors doesn't approach because of too large number of data [9]. One of such scenarios carries out an important role. It normalizes the quality of the document's electronic copy. This algorithm allows improving the quality of the scanned materials significantly [10]. Usually, after scanning the background of the document turns out to be not white, but grayish partially or completely (Figure-1), and on the majority of images there are artifacts of black strips received at the edges and at the places of the book's excess [11].

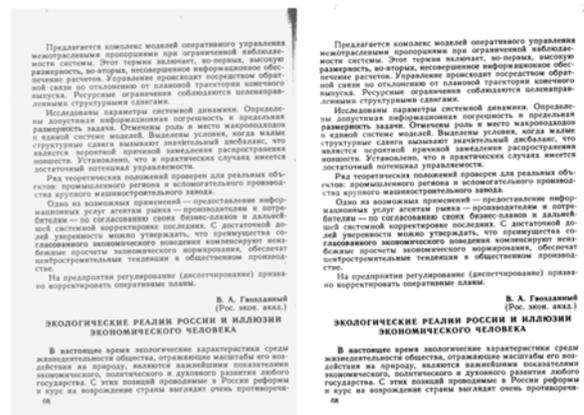


**Figure-1.** A fragment of the scanned image showing the quality of the lines/the text and the artifacts of black stripes in the places of book's excess.

#### 2.6. The check of the compliance to the set criteria

It was established experimentally that if to reduce the brightness, the quality of the lines/the text spoils and if to keep them in a monochrome quality, there will be rigid transitions reminding ladder steps.

The next step was a need of increasing of the contrast of images with a replacement of all excess colors at their monochrome analogs, but within the set range became the following step, thus, the problem of the improvement of the quality of initial materials was successfully solved (Figure-2).



**Figure-2.** The original image (from the left side) and the image after the transformation (from the right side).

The algorithms of a formation of simple and difficult inquiries, with a possibility of data's filtration which are in the database on one or several criteria were also developed and introduced: the author(s), the name of the work, the place and the year of the edition, etc. The algorithm of the search in any field of the bibliographic description of a source was created [5].

### 3. RESULTS AND DISCUSSIONS

Within the practical researches the following aims are realized:



- the portal interface is developed , including a mobile version;
- the algorithms of a search according to the set of criteria of a database are developed: the author, the name, the publishing house, etc., for an ensuring of the flexible search;
- the algorithm of access' management on the basis of the security policy accepted at University is introduced;
- the portal of full text electronic separate collections of the books and magazines is created on the basis of library stocks of REU;
- the hierarchical system of catalogs (collections of documents) is created;
- the problems of a visualization of reading process adapted to the various devices (the smartphone, the tablet, the computer) are solved;
- the tool kit for a filling and editing a database is created.

At the moment the REU hypertext electronic and digital scientific library of G. V. Plekhanov passed the stages of primary programming and control, it is introduced and is tested actively. The modern hypertext provides the organization and submission of information, and its wide circulation is connected with that this technology can provide more effective and personalized access to the information streams, based on the communications both inside, and between the documents [7].

#### 4. CONCLUSIONS

In the conclusion we will note that despite all the difficulties connected with the creation, the testing, the control and the completion to a condition of the ready decision electronic hypertext library there can be a universal remedy of representation, accumulation and transfer of knowledge both in the form of traditional materials, images and texts and interactive presentations, an audio and video in the systematized look available to search and the subsequent demonstration that certainly expands possibilities of researchers.

#### REFERENCES

- [1] Alekseeva T.V. 2012. Obuchenie studentov s ispol'zovaniem oblachnykh tekhnologii, Slavyanskii forum. 1(1): 81-87.
- [2] Bortsova D.E., Nedelkin A.A., Olenov L.A. 2015. Osnovnye etapy i napravleniya razvitiya sovremennogo giperteksta. Slavyanskii forum. 2(8): 50-59.
- [3] Lutovinova O.V. 2009. Gipertekst: ponyatie, osnovnye kharakteristiki, vozmozhnye podkhody k lingvisticheskomu analizu, Izvestiya Volgogradskogo gosudarstvennogo pedagogicheskogo universiteta. 5: 4-7.
- [4] Zharkova U.A. 2015. Diskursivnye kharakteristiki web-saita universiteta: k probleme strukturovaniya kontenta. Filologicheskie nauki. Voprosy teorii i praktiki. 5-2(47): 59-61.
- [5] Antonenkova A. V., Bortsova D. E., Stepanova M. G. 2015. Sovremennye podkhody k avtomatizatsii protsessov byudzhetrovaniya kommercheskoi organizatsii. Lizing. Tekhnologii biznesa, 8.
- [6] Titov V.A., Nedelkin A.A. 2014. Gipertekstovaya elektronno-tsifrovaya nauchnaya biblioteka, Mezhdunarodnyi zhurnal prikladnykh i fundamental'nykh issledovaniy. 12-1: 135.
- [7] Romanova Yu.D., Dyakonova L.P. 2013. Sovremennye obrazovatel'nye standarty i sertifikatsiya v oblasti informatsionnykh tekhnologii. Perspektivy razvitiya informatsionnykh tekhnologii. 16: 114-118.
- [8] Gvozdeva V.A. 2013. Informatika, avtomatizirovannye informatsionnye tekhnologii i sistemy. ID FORUM, NITs INFRA-M, Moscow, Russia.
- [9] Zhelondievskaya L.V. 2011. Gipertekst graficheskogo dizaina. Vestnik Orenburgskogo gosudarstvennogo universiteta. 9(128): 54-57.
- [10] Potapov A. A., Pakhomov A. A., Nikitin S. A., Gulyaev Yu. V. 2008. Noveishie metody obrabotki izobrazhenii. Fizmatlit, Moscow, Russia.
- [11] Stepanenko O. S. 2005. Skanery i skanirovanie. Kratkoe rukovodstvo. Moscow, Dialektika.
- [12] Aleshin L.I. 2011. Informatsionnye tekhnologii: Uchebnoe posobie. Market DS, Moscow.
- [13] Goncharenko L.P., Yakushev A.Zh., Sybachin S.A., Romanova Yu.D., Miloradov K.A., Antonenkova A.V., Titov V.A. 2014. Avtomatizirovannaya portal'naya gipertekstovaya elektronnyaya sistema dlya provedeniya nauchnykh meropriyatii v формате konferentsii i konkursov. Khroniki obedinnogo



fonda elektronnykh resursov Nauka i obrazovanie.  
12(67), vol. 1, p. 140.

- [14] Ibragimov I.M. 2008. Informatsionnye tekhnologii i sredstva distantsionnogo obucheniya: Uchebnoe posobie dlya stud. vyssh. ucheb. Zavedenii, ITs Akademiya, Moscow, Russia.
- [15] Ivanova E.A. 2005. Kontseptsiya giperteksta, Informatsionnye tekhnologii i prepodavanie inostrannykh yazykov. MESI, Moscow, Russia.