

E-SERVICES: COLECTIVE MEMORY, SETTING UP NEWSPAPER DIGITALIZED ARCHIVE CASE STUDIES

Marius Stoianovici, Ionela Bîrsan, Ana Maria Cornelia, "Transilvania" University of Braşov

Abstract: Digitalizing comprise all life cycle stages of a digital document: digitalizing means the conversion of the document from its original analogical support, to digital format, objects and documentation references description and representation, processing, access provision and long term preservation of digitalized content.¹ Information and communication technologies are contributing to research in any area of knowledge, while digitalizing can be perceived as „more than a viewed page substitute, it can generate also new research opportunities.”² Text representations can send to other related texts, to different approaches of other authors. Those which are bringing the latest information are the periodicals, which are increasingly available in electronically format, either through digitalization, or available only in the digital media.

The means of making available to user the paper printed documents, through digitalization, are including photography or scanning (in either image or text versions). The digitalization and digitalized archive of Monitorul de Brasov Collection within Transilvania University of Brasov outcomes are presented. The metadata are data about data and are aimed at identifying and describing any digital resource. Special collection digitalization projects should provide an „information critical mass” user targeted but first the user type identity has to be defined. A complete guidebook for developing a digitalizing project by metadata creation, metadata charts, interoperability protocols and online dissemination of digitalized documents is achieved.

Keywords: digitalizing, metadata, interoperability, digital archive, Koha.

Introduction

Digitized archive developed through various digitized platforms models is widely used by info documenting agencies. Considering the printing company initially set to issue printed newspapers and detaining classical newspapers collections the digitized archive concept is less common. Digitizing these documents and their subsequent upload to an appropriate platform is the solution for their preservation. The authors found an optimal solution of the digitize-preserve- automate retrieve process of for the classical paper documents. The DIGIBOOK 2NET scanner is used to digitize the documents. The Koha integrated library system is used for data recording and storage of each digitized document. An online catalogue is created and offered to the newspapers readers as a mean of data retrieving. The Koha system generates automatically barcodes for each of the digitized document. These barcodes are retrieved by a document identifying scanner. This system allows an inventory of the classical archives.

Retrieving information integrated open source systems

Integrated library management systems can cut costs and enhance the efficiency of library services and therefore are absolutely necessary for the management of housekeeping operations.[7] But small and medium- size libraries face a difficult situation due to the high cost of commercial library management systems, available in the market.

For at least the last two decades, libraries have overwhelmingly obtained their core automation systems from specialized vendors who offer the software through licenses that allow the company to retain exclusive access to the underlying source code. In recent years,

open source software has become an increasingly popular alternative. The underlying program code is made available for anyone to inspect, repair, or improve. The open source software movement has entered the library automation industry, introducing a new set of integrated library systems and a clique of companies offering a business model based on service and support rather than software license fees. [1]

The library integrated systems that are used in the Romanian university libraries are software products that were acquired from different suppliers. The investment to digitize the department library which is being developed -but is not designed to exceed 10,000 volumes in the near future- would be useless at this stage considering the availability of the open sources library integrated systems that are used worldwide.

Naturally libraries without much/appropriate financial resources need a cost effective way to automate their services. Free/Open source software was a revolutionary concept among computer programmers and users. To a certain extent free/ open source solutions could provide an alternative to costly commercial application software's. Free/Open source software offers to its users the ability to run, copy, distribute, study, change, share and improve for any purpose. [6]

Koha- library integrated management system

Koha is an integrated library management system that was originally developed by Katipo Communications Limited of Wellington, New Zealand for the Horowhenua Library Trust (HLT), a regional library system located in Levin, some 100 kilometres north of Wellington. Libraries considering implementing Koha have an option to hire Katipo staff to help with the implementation, and there is also a list of other organisations who could be hired at <http://koha.org/installation/support.html>. [5]

The Koha project uses a number of channels to allow members of its community to communicate with each other — there is a general mailing list, as well as separate ones for developers, Windows users, French-speaking Koha users/developers, and German-speaking Koha users/developers. In addition, the developers use Internet Relay Chat (IRC), a real-time message facility for scheduled meetings and less formal conversations. [3]

Koha 3.0 was selected because the GNU licence (open source) was considered more future-proof than proprietary products, and more open to customisation to meet the special needs of the library. [2]

Thanks to the efforts of the open-source community, any library can now enjoy a serious ILS at no licence cost: the saved money can hence be allocated to the extension of the collection.

Koha satisfies all the functional requirements of a library management system. In addition to the functional modules(acquisition, cataloguing, OPAC, circulation, and serial control), Koha can provide some features that are only available with costly ILS; web OPAC, document status inquiry, reservation and holds of documents through OPAC, customization of user and graphical interface, import and export of MARC data etc. We implemented Koha, library integrated system. Koha has all modules needs to create online catalogue of digital archive.

We can choose the type of barcode. (Fig.1, Fig.2, Fig.3)

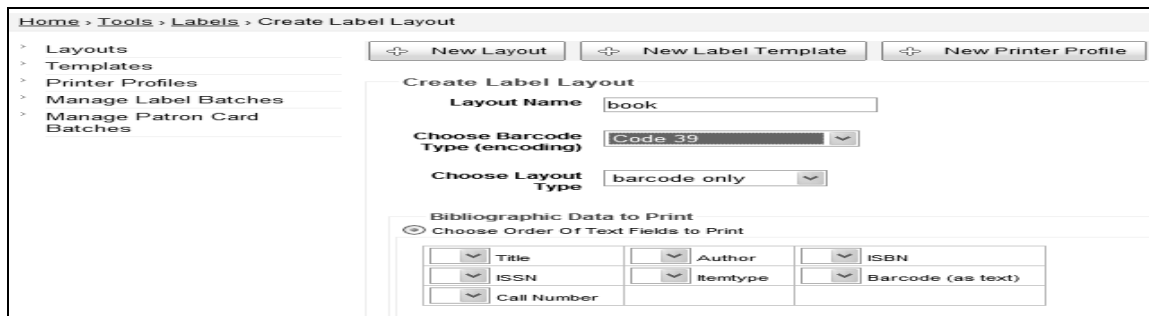


Fig. 1 Create label layout for Koha barcode

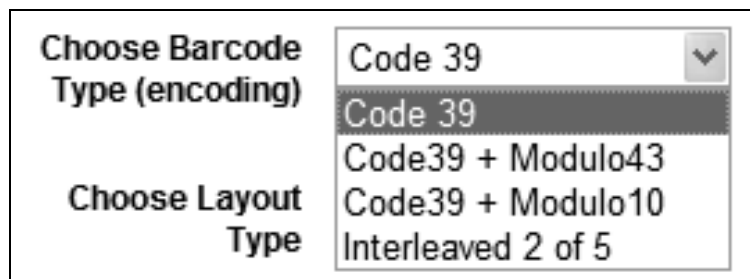


Fig. 2 Choosing layout type for Barcode



Fig. 3 Barcode for printing

We choose a layout type from the "Choose Layout Type" dropdown menu. From every barcode we choose digit representation.

Digitalizing equipment

Book2net (Figure 3) production is a specific book fast digitalizing development. Due to its 2000 page /h scanning speed Book2net offers an unmatched productivity. It comes equipped with:

- 40 or 50 mega pixels sensor
- Hi-tech sensor

- Non-reflection Makrolon auxiliary/supplementary board
- It works with an external PC; scanner comes with Book2net expert software offering a optimized post processing flow .



Fig. 4 Scanner image

Facilities:

- **Sensor's technology:** book2net flexibility; Possibility to choose between sensors resolution of 10.5, 32, 40, or 50 megapixels meeting very particular needs.
- **Only 0,3 seconds/ scanning: 0,3 seconds** impressive scanning speed and a 1,9 seconds processing cycle offering an no matching productivity
- **A2 book support:** up to 10 cm automatic lifting; non-reflective, low abrasive worktable overlay
- **UV and IR radiation free:** book2net light source is exposure radiation free
- **focusing depth 12.5 cm**
- **Formats:** jpg, jpg 2000, tiff and pdf simple or multiple etc.

Turning on: plug the unit to mains and put the main switch to On.

Then the book expert soft is opened and customized along as main scanning stages are operated.[4]

Initial calibration is required in case of any redeployment or lighting variations in the working area and this has to be performed before any other procedures in order to obtain accurate scanning. (Figure 5)

After calibration, using “settings- customize scan process” command, the particular scanning elements as: dividing for two pages scanning area, color settings and resolution setting, but not only, can be set.

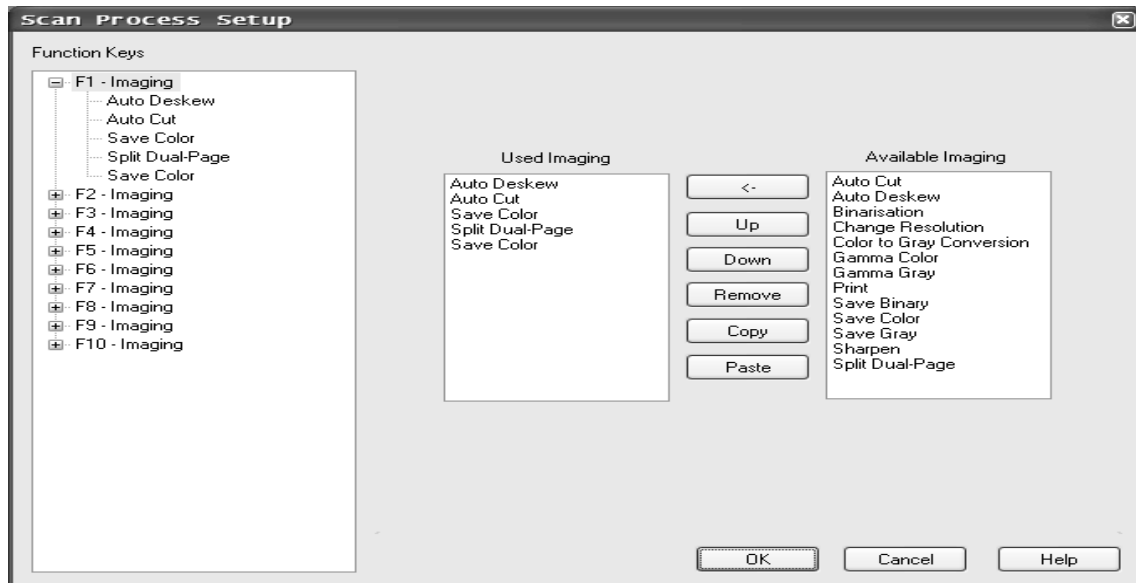


Fig. 5 Image of book expert soft

Next, by using “settings- customize scan job- scan” command, the actual working area is selected, as shown in the figure bellow, in order to optimize scanning duration as well as setting the limits of the scanning area.

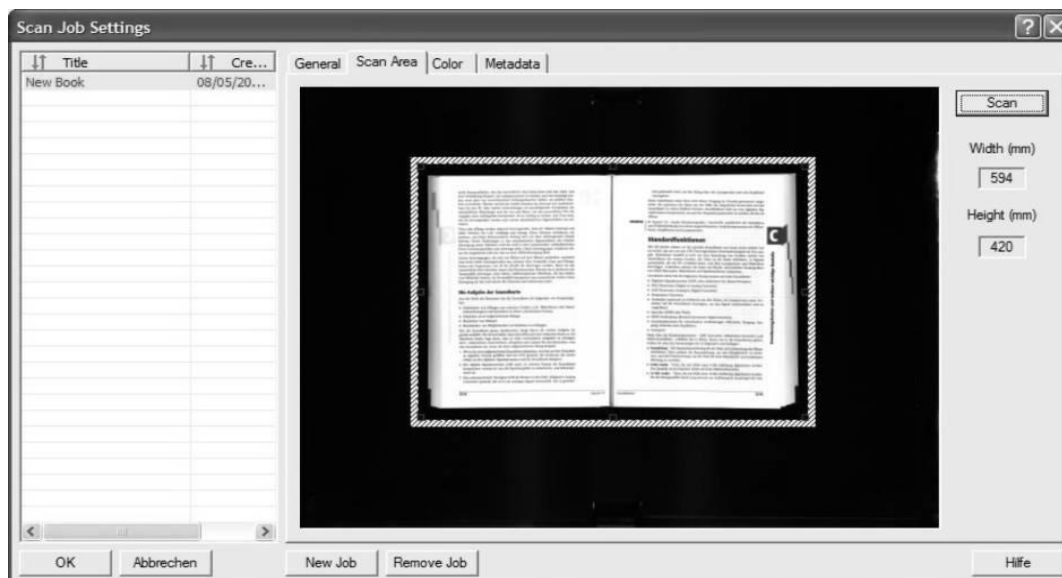


Fig.6 Limited of scanning area

After these stages are covered –setting the above parameters included- one can proceed to actual scanning.(Figure 6)

There are two ways to activate the scanning command: by foot command from the ground pedal or by hand, at the worktable level.

A post-processing scanning output images stage is required aiming to:

- ✓ manually delete the excess scanned material;
- ✓ image rotation if its position has altered during scanning

- ✓ performing some adjustments consistent with the scanning appearance/visual aspect

Scanning stages

Initial stage consists in starting the equipment and operating soft as above described. Afterwords the sensor has to be calibrated to the specific environment circumstances. The document is placed on the worktable and using the bookexpert operating system the area of interest is set, as presented in figures 7 and 8, in order to optimise the acquiring process.

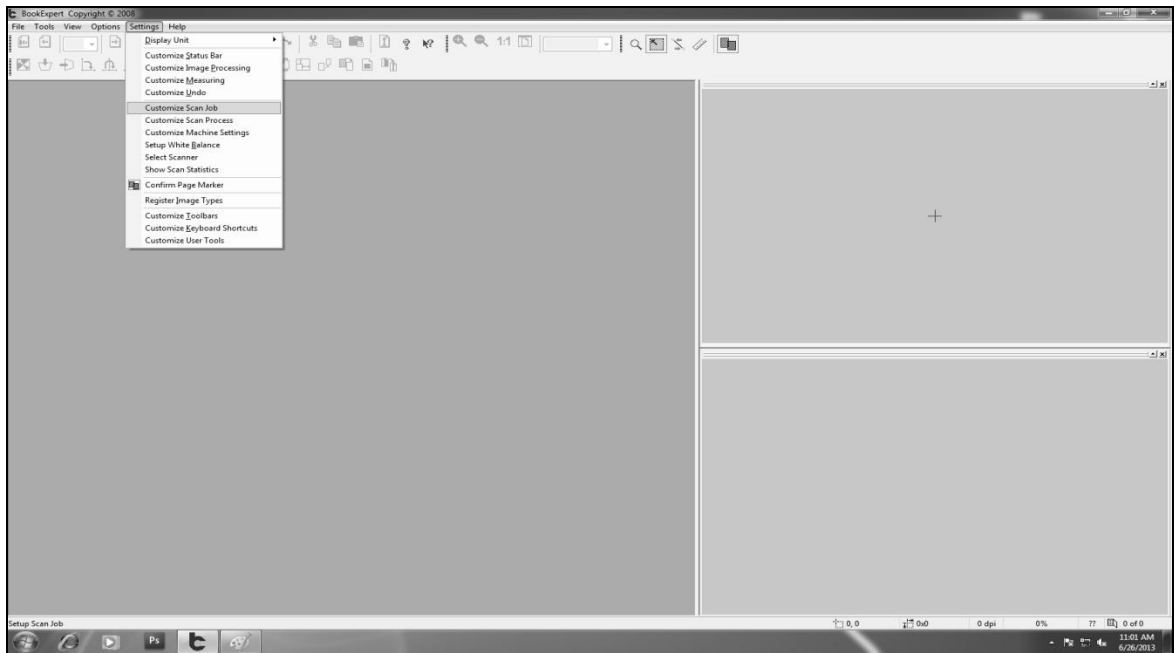


Fig.7 Operating system set I

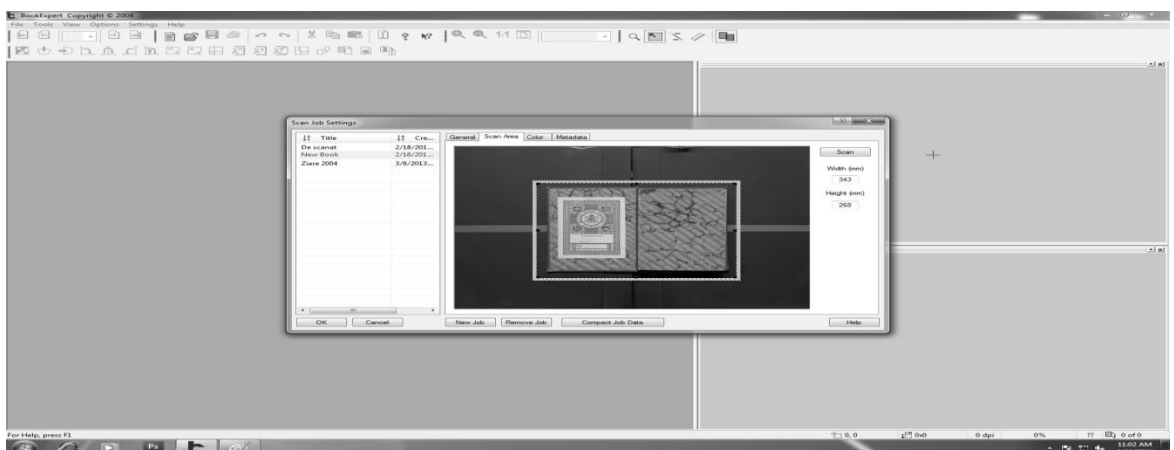


Fig. 8 Operating system set II

Accessing settings- “customize scan process” command (figure 8) the scanning customization is performed, introducing certain elements such as: dividing for two pages the scanning area, color settings, resolution setting and so forth, according to particular scanning requirements.

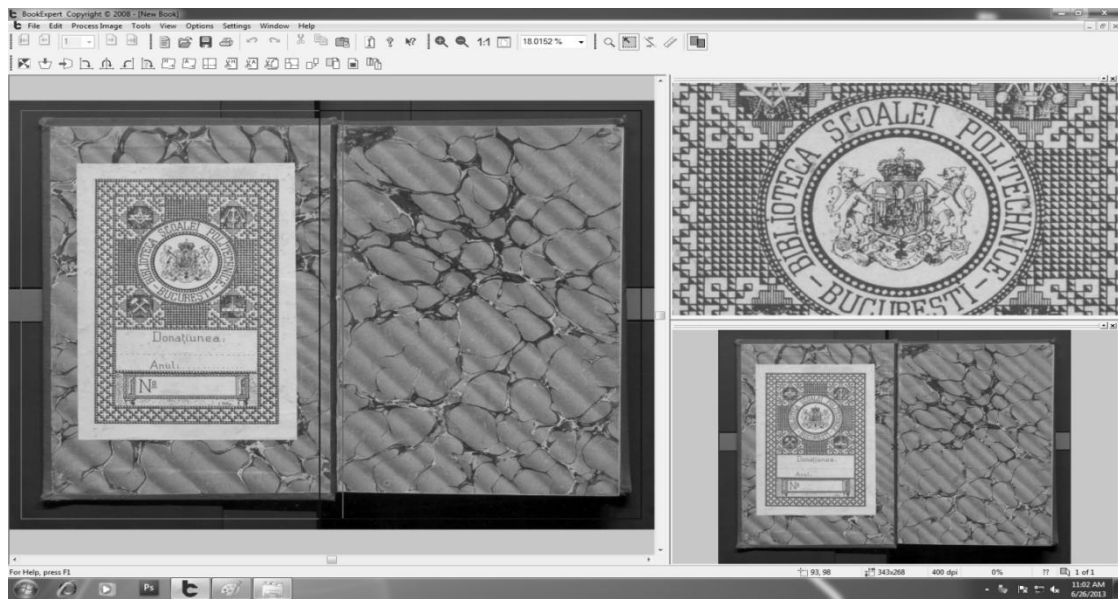


Figure 9 Raw scanned image

Actual scanning starts by pressing the foot pedal. Raw scanned image is presented in figure 9. As seen there are two different color frames, red and blue, helping to divide the work surface in two areas corresponding to the two book pages. The frames adjustment is presented in figure 10 and it will be performed taking in to account possible book movements during scanning..

In the next scanned pages the proper framing is seen (figure 11). The excess scanning will be manually deleted later, during the image finishing procedure

After the entire book is scanned the default created folder- by the software, during the stage presented in figure 8- can be accessed.

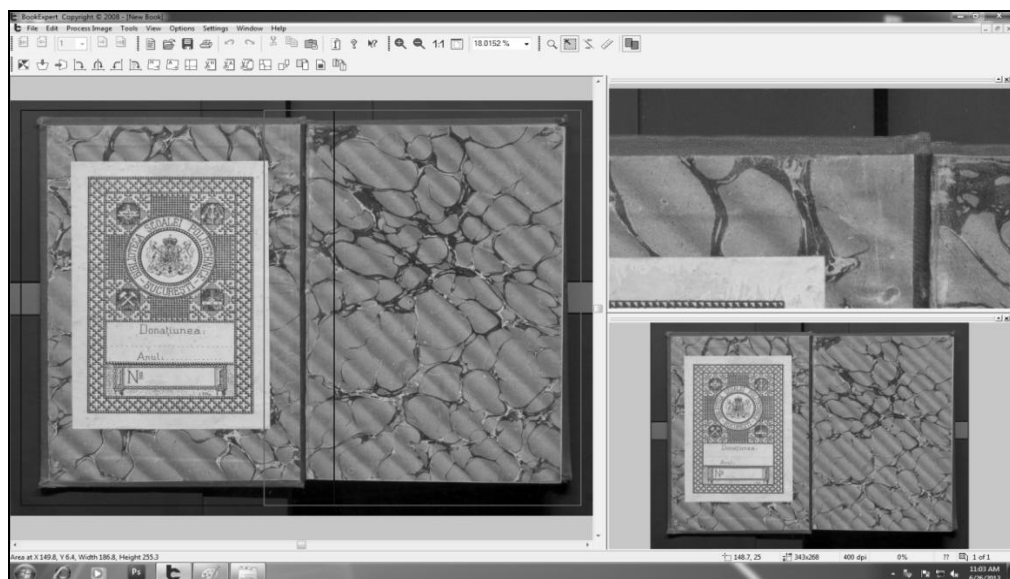


Fig.10 The frames adjustment

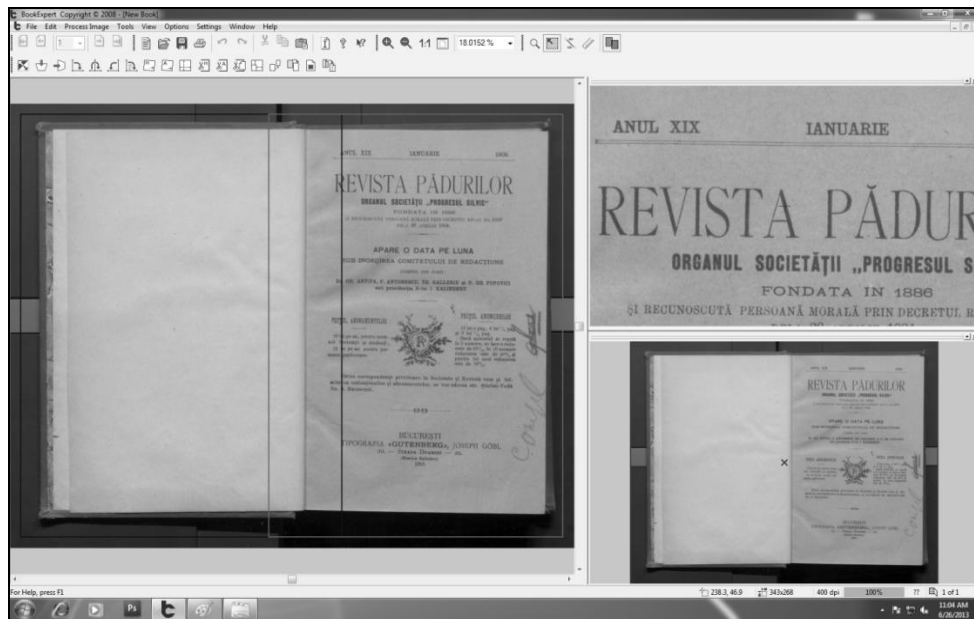


Fig.11 The proper framing

From this folder a set of images can be imported (figure 12 a) and processed using book expert software specific tools. The main tool is offered by the “manual cut” command, figure 12 b. The desired area is selected and using the above mentioned command and the useless information is deleted, as in figure 11 b, the outcome is the final image (figure 13). If some rotations have occurred these can be sorted using the “rotate” command placed in the same command bar as “manual cut”.

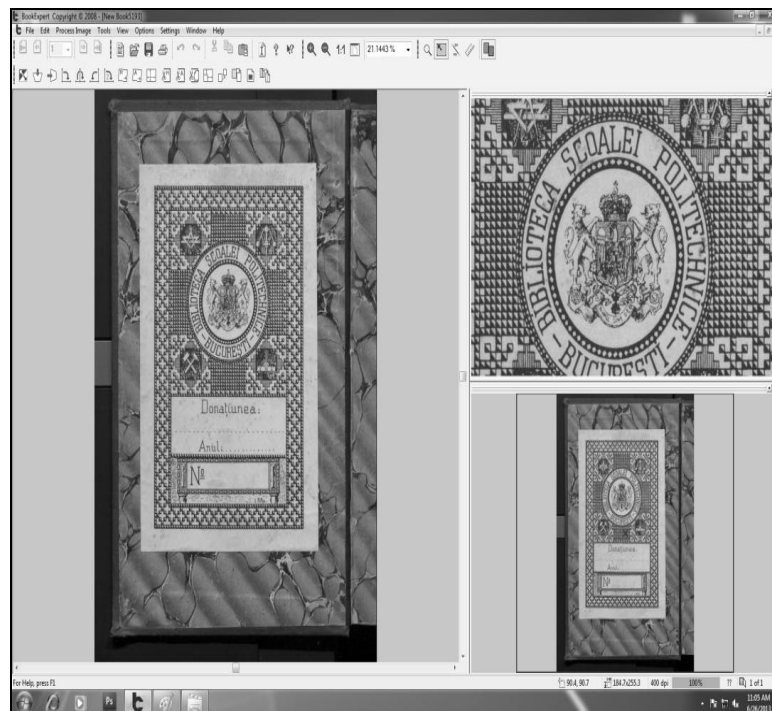


Fig. 12 a) Images imported b) Manual cut

Thus for each raw image a correspondent finished image will be obtained. In this case it has the following properties: tiff format with a 400DPI(Dots Per Inch) resolution. Depending on the ultimate goal the resolution can be altered. Therefore for an internet visualization 75 DPI resolution is used while for a quality printing 300 DPI resolutions is required. The resolution will influence the quality and the quantity but also the maneuverability of the stored image/information.

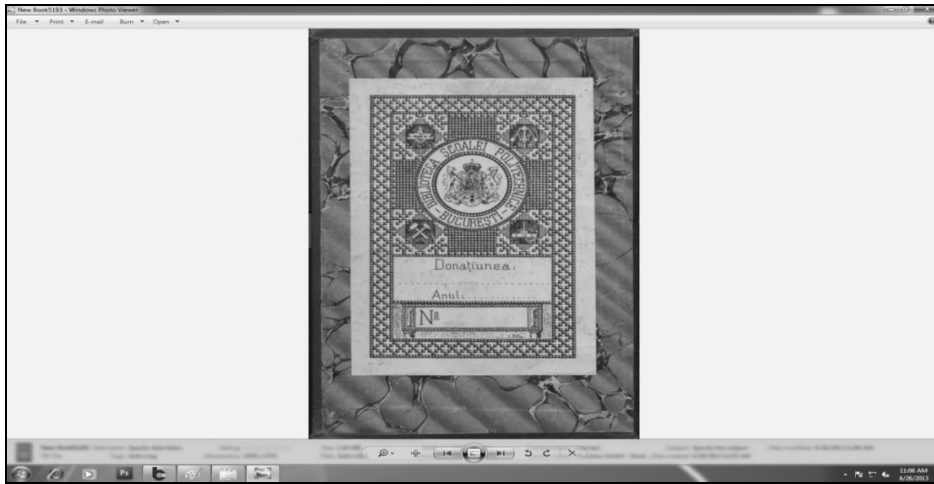


Fig. 13 Final image I

Scanning itself is performed in a short amount of time which depends both on all the initial settings as well as the operator's experience.

Conclusions

The use of this procedure is a particularly efficient modality of preserving the information, and of enhancing the users' access to information.

We deem this particularly modern method may ensure performing a double mission: preserving of valuable documents and full access to original information.

The proposed solution offers a minimum cost model using open sources software and it offers also a useful instrument not only in higher education, but also available and recommended for museums, research institutes, media industry.

Acknowledgement: This paper is supported by the Sectoral Operational Programme Human Resources Development (SOP HRD), ID134378 and ID 137070 financed from the European Social Fund and by the Romanian Government.

References

- [1] American Library Association. (2008). Chapter 3: Major Open Source ILS Products. *Information Science and Library Science* , 44 (8), 16-31.
- [2] Bissels, G. (2008). Implementation of an open source library management system. Experiences with Koha 3.0 at the Roya London hoeopathic hospital, Program:electronic library and information systems. *Vol.42 Iss.3* , 303-314.

- [3] Chawner, B. (2004). *School of Information Management, Victoria University of Wellington*. Retrieved March 15, 2011, from Free/Open Source Software:New Opportunities, New Challenges: <http://vala.org.au/vala2004/2004pdfs/33Chawn.PDF>
- [4] i2S-DIGIBOOK – http://www.i2s-bookscanner.com/pdf/digibook_mag_no6_en.pdf
- [5] Koha, L. &. (2009). *The Koha -- Open Source Library System*. Retrieved March 12, 2011, from Koha: <http://koha.org/documentation/manual/3.2/tools/label-creator>
- [6] Vimal Kumar, V. (2005). Free/Open source integrated library management systems:comparative analysis of Koha, PHPMyLibrary and OpenBiblio. *Modern trends in IT application in Library and Information Servises*. University of Calicut.
- [7] Vision, T. (2011). Retrieved March 15, 2011, from <http://www.tagvision.dk/products/akmuk.php>