

Digitalization models and algorithms for preservation of Armenian rare books

Tigran Zargaryan
Fundamental Scientific Library of the
National Academy of Sciences of Armenia
Yerevan, Armenia, tigran@flib.sci.am

ABSTRACT

Preservation conditions for the rare books being kept in the Armenian libraries are not satisfactory. Many rare books which were indicated as sources in the publications of scholars at the beginning of the twentieth century are currently unavailable as they have been destroyed, either through poor storage, or because of the fragility of the paper or lost in war and civil strife, and there are no surrogate copies for them. The same will happen with the remaining ones, as the paper continues to deteriorate.

This paper provides with the models and solutions being in use in the Fundamental Scientific Library for preservation purposes.

Keywords

Information technologies in the libraries, preservation.

The first Armenian book was printed in Venice in 1512 by Yakob Meghapart. Between 1512 and 1513 he printed five titles: "Urbatagirk", "Parzaytumar", "Pataragatetr", "Altark" and "Tagharan". Following the initial period Armenian printing spread rapidly during the XVI-XVIII centuries into cities with large Armenian presence such as Venice, Constantinople, Rome, Lvov, Milan, Isfahan, Livorno, Marseilles, Amsterdam, Madras, Calcutta, Smyrna, Etchmiadsin, Trieste, Petersburg, Nor Nakhijevan, Astrakhan. Thanks to Armenian bibliographers Dr. Ninel Voskanyan, Revd. Dr. Vrej Nersessian and other scholars the early printed book collection has been catalogued to the highest standard in three sequences: 1512-1800; 1801-1850 and 1851-1900. In European libraries the largest Armenian collections are held in the British library, followed by the Bodleian library Oxford and BN in Paris.

The Fundamental Scientific Library of the National Academy of Sciences (NAS) has one of the largest collections of rare books and periodicals in Armenia: 13,540 rare books in

Armenian, English, French, and German, Latin, Slavonic languages. The conservation status of these collections is poor. The fluctuation of temperatures and level of humidity in the stacks during the autumn and spring seasons and the pollution level remain uncontrolled. All collections are very fragile, suffering from paper deterioration and fungus lesion, and the physical condition of the books is rapidly deteriorating. In the library readers have access to the originals, and in some instances the only copy of some titles can be found in the FSL collection. Digitized copies will allow the library staff to make available the surrogate copies and focus on the care and conservation of the originals. The European language material in the FSL is of crucial importance for the study of Armenia and its ties with the countries and cultures of Asia Minor in the Middle Ages. Most of these are still unavailable to users, because of our need to take care of the material and prevent further damage. For instance in the atlases 'Atlas Minor: Totius Orbis Terrarum' and 'Atlas Nouveau: Contenant Toutes Les parties Du Monde' (publication year 1696) a researcher can find lot of maps of the countries, regions (including Asia Minor) in Medieval times. In the book 'Voyages de Corneille Le Bru: par la Moscovie, en Perse' (publication year 1718) special chapters (with images) are about the Armenians and their neighbours, their history, habits, customs and traditions. All these books need urgent digitisation.

Having these in minds, the library decided to: (i) establish a digitization Centre in the Fundamental Scientific Library; (ii) implement a modern scanning and conservation technologies that will match the best in European Libraries and standards for the care of vulnerable resources; (iii) create a rare book and journal digitized e-library with

metadata structured records compatible to international standards, and to make this digital collections accessible to researchers via a Web; (iv) create from the original materials a high quality master copy (TIFF v 6.0 uncompressed) and an access copy (in jpeg or PDF formats), link them with the relevant bibliographic records from the Armenian Libraries Union Catalogue¹; (v) train a team of personnel for future scanning projects; (vi) start optical character recognition process for these collections (where applicable).

Such as the library was not having any professional digitizing equipment, and the core element of our project was presence of high resolution photo camera, as a first step we started to explore the list of possible donors, supporting preservation activities for the rare books and documents. As a possible candidate we stopped our decision on EAP. The aim of the Endangered Archives Programme² (EAP), supported by Arcadia, and managed by the British Library is to contribute to the preservation of archival material that is in danger of destruction, neglect or physical deterioration world-wide. The main means by which the Programme achieves this is through the creation of digital or microfilm copies of endangered materials and the relocation of the originals to a safe local archival home.

The library prepared a grant proposal and applied to the Programme. The application was considered by an EAP International Advisory Panel and was recommended for funding.

The digitizing system consisting of Phase One P45 Digital Camera System, ICAM GUARDIAN Book Handling System, 1000VA UPS, Graphic Workstation Computer with quality 19" monitor and 2 x 500GB external hard drives, Network Attached Storage RAID System with 2TB fast access memory was obtained from Icam Archive Systems LTD³. Currently digitization of Armenian rare books and Armenian periodicals is in process.

¹ <http://www.armunicat.am:8991/ALEPH>

² <http://www.bl.uk/about/policies/endangeredarch/homepage.html>

³ <http://www.icamarchive.co.uk>

Software solutions. One of the goals of the preservation project is to mount images of the rare books on the Web and make them accessible to the researchers. As a perspective direction for the library and archives automation and digitization FSL is orienting its activities on using Free/Open Source Software (FOSS) solutions. For example, **ePrints** package is in use for preservation of journal collections published in the National Academy of Sciences⁴. For mounting on the Web the images of rare books we are using **Greenstone** content management system, which is again a FOSS product. The collection of the FSL rare books is accessible at: <http://www.flib.sci.am/eng/?q=node/5> . Visit the link 'Armenian rare Books'.

Further plans. Digitizing activities are already started in FSL and we are thinking about our future plans. We see 3 main directions, where we must concentrate our efforts.

- **Connecting digital collections to the Computing Grid for South Caucasus Region** using ArmCluster resources. Institute for Informatics and Automation Problems of NAS is managing the system, and currently the library is collaborating with IIAP for possible solutions. Successful realization of this task will allow to make the process of building digital collections much faster, and through the main European multi-gigabit computer network for research and education purposes **Géant** to integrate the collections to the "Memory of the World" database.
- **Optical Character Recognition of images.** Our experiments with the images of Armenian rare books and periodicals demonstrated that none of existing OCR packages is able to recognize Armenian scripts which have been in use in XVI-XIX centuries. This means that we must develop our own OCR product (probably open source), which will be able to solve our problems on

⁴ Visit the section e-Publications at <http://www.flib.sci.am/eng/?q=node/5>

character recognition. Here we seek support and advice from mathematicians and software developers.

- Focusing our activities on using UNL as a tool for translating of the bibliographic records from the Armenian Libraries Union Catalogue and full text databases to the UNL language.