AN EXPERIENCE OF PRESERVATION AND EVALUATION OF DOCUMENTARY HERITAGE THROUGH DIGITIZATION: "MANUSCRIPTORIUM" PROJECT

Elena Tîrziman, elena.tirziman@bibnat.ro National Library of Romania

Abstract

Within the framework of Manuscriptorim project coordinated by the National Library of the Czech Republic, the National Library of Romania started the program of digitizing the most valuable collections of old books and manuscripts.

The paper details the methods and techniques used between the two libraries to implement the project, creation of the metadata and description of the books and manuscripts selected. The Romanian contribution to the project reflects the main stages of the evolution of printing between XVI and XVIII century in the historical provinces inhabited by Romanians.

Keywords

Manuscriptorium, ENRICH, Digital content, Digital library, Romanian old books, Manuscripts, Batthyaneum

Introduction

"Manuscriptorium", the project coordinated by the National Library of the Czech Republic, aims to build a virtual research environment for historical resources. Manuscriptorium began by making available on the Internet and by collecting information on book type of historical resources connected to a virtual library of digitized documents.

The Manuscriptorium project has its origins in UNESCO's initiative "Memory of the World" from 1992. Czech Republic participated in this program with the project "Memoriae Mundi Series Bohemica", the project which allowed the

identification and representation of old books in the electronic context (www.unesco.org).

Since 2002, the accumulated experience allowed the further development of the Manuscriptorium project by creation of a virtual library that facilitated the Internet access to old and rare books, manuscripts as well as an electronic context for information, documentation and research. The main contributors to this project are libraries, museums, archives and other cultural institutions holding such cultural heritage. The Manuscriptorium platform has allowed the exploitation of digital content through other specific projects such as ENRICH (European Networking Resources and Information Concerning Cultural Heritage) and REDISCOVER (Reunion of Dispersed Content: Virtual Evaluation and Reconstruction)

The objectives of the project

The objective of Manuscriptorium (www. manuscriptorium.com) is to provide online access to what cultural written heritage means, especially old books and manuscripts but also other documents from the category of the archival documents (historical documents, administrative documents, maps, etc.) as a secondary priority. All digital documents in Manuscriptorium represent a copy of an original physical document existing in a collection or a copy of a document that no longer exists in paper format (the digital document is realized from a copy of the original document).

The secondary objectives of the project are related to the preservation of the documents in digital library through their digitization and their use in multiple forms, from simple catalogs to multinational virtual exhibitions.

Like any other library, the Manuscriptorium digital library also has an online catalog and provides access to full text documents. The catalog of these resources is very accessible for the user and provides detailed information about each historical document included, indicates where the physical document may be found and also provides an electronic copy for many documents. In the first stage of Manuscriptorium's existence, many of the documents included in this digital library were an alternative to the study in the reading rooms for consulting the original physical documents or the copies in Microform.

Now, the access to the full text resources through the Internet, in many cases eliminated the reading rooms and also gave the possibility to access geographically distributed resources, in other words, it provides access to digital resources in the collections of various libraries, museums, archives or private collections in different countries or geographical areas.

Methods and Techniques used

Manuscriptorium development cooperation principles have crystallized together with the evolution of this digital library and resulted in a technical contract and a legal contract of participation. They express the collaboration agreement in a technical aspect and from the dismissal and use of digital content between the National Library of the Czech Republic responsible for the coordination, management and development of this complex digital product and partner institutions contributing to digital contents.

The technical contract is based on the real possibility of partners to get involved in supporting and developing their own collaborative digitization projects and digital library services. From a technical point of view, the Manuscriptorium involves two categories of partners:

• Partners who have developed projects and digital library collections and have the IT infrastructure which enables communication, protocol interoperability through OAI-PMH (Open Archives Initiative Protocol for Metadata

Harvesting);
- Partners who are unable to ensure interoperability through OAI-PMH protocol but who have digital collections or at least digitized documents.

For each category of partners, the methodology is different; for the first type, the gathering of information and its integration into the digital library is achieved by harvesting; the other uses the M-TOOL application that allows making of the necessary descriptions for digitized documents.

The documents description format was permanently improved according to the research level. Originally, the SGML format was used and then the HTML format to enable viewing of documents by any known browser without the need for additional adjustments by the user. The emergence of XML format and particularly the version MASTER DTD especially for the description of manuscripts, has led to the usage of this format.

The image formats supported by Manuscriptorium are the formats recommended by the Web consortium: JPEG, GIF or PNG.

The criteria that a document must meet to be included in Manuscriptorium are:
- The value and significance of the document reported to the objectives of the Manuscriptorium digital library;
- The usefulness of the information about the document from a technical point of view (metadata) and the information contained in the document (full text, image, sound if any);
- Quality and reliability of the data.

The National Library of the Czech Republic is the one who takes the final decision whether to include a document in the Manuscriptorium system.

The Manuscriptorium platform was designed in its entirety to be a comfortable and useful information service for users. They had taken into account the possibility of passing from one to another historical document, procedures for access and communication, the use of a unitary interface regardless of which computer requires access and where it is located.

Beyond the possibility to search in disparate collections found in different institutions and facilities of using the retrieved images, the added value of the Manuscriptorium digital library is also given by the provided retrieval, evaluation and usage tools and applications. An example is the support in retrieving useful information and documents in the multilingual context created in Manuscriptorium. The difficulty is due not only to many languages used and retrieved in this digital library, but also to the fact that there are languages used in certain historical periods, in certain geographical areas, written in different alphabets, even expressing dialects or grammatical issues not used today. These are documents in Latin, Slavonic, Greek, Arabic, Indian alphabets from different historical stages and from different geographical areas. Therefore, to achieve a unified and integrated retrieval application, many linguists and specialists needed to work on different aspects of studying these documents and also it was necessary to use local applications of libraries holding such documents.

Another feature available to users is the ability to customize the usage of Manuscriptorium, to create a user profile and to realize their own virtual library according to their research. This information and documentation mechanism works in two ways: static and dynamic. Static means that the addition and deletion of documents is made by the user when searching. Dynamic - these operations are automatically realized when the search demand is defined.

ENRICH Project

A higher stage in Manuscriptorium's evolution is the project ENRICH - European Networking Resources and Information Concerning Cultural Heritage (Uhlir and Knoll, 2009), a project developed within the frames of the eContentplus program. Also coordinated by the National Library of the Czech Republic, ENRICH aims at virtual aggregation of information regarding old heritage documents, especially manuscripts and incunabula and other historical documents. The main purpose is to provide free access to the old heritage documents from various European cultural institutions in order to create a shared virtual research space.

ENRICH is built on the platform of the digital library Manuscriptorium, which on the date of the submission of the proposal already had managed to aggregate and manage data from 46 collections in the Czech Republic and abroad. ENRICH was conducted between 1 December 2007 and /November 30, 2009.

ENRICH partners are European institutions holding cultural heritage fit for the purpose of the project: national libraries, university libraries, archives, museums and culture institutions.

ENRICH consortium consisting of 18 partners was supported by a large number of other institutions holding important documentary collections.

The intention to cooperate in the project development was expressed, among others, by the national libraries of Hungary, Kazakhstan, Moldova, Poland, Romania, Serbia, Sweden, Turkey and the university libraries in Bratislava and Heidelberg.

Through ENRICH, Manuscriptorium develops and expands its technological and informatics services facet, expands its users' typology assimilating together with researchers and students, the academic and cultural environment and even the general public; it diversifies its products and services typology (including personalized services); it contributes to the development of theoretical and applied research in information and communication sciences and computer sciences; integrates digital collections from a wide geographical area.

Romanian contribution to the project

The National Library of Romania is a partner in the project Manuscriptorium / ENRICH since April 2008 (Tîrziman, 2008) and contributed with 109 Romanian books from the XVIth and XVIIth century, with a special cultural, historical and artistic value. Most of them are religious works, but there are also law and history works. We can mention a few names of scholars and printers such as Mitropolit Dosoftei, Varlaam, Macarie, Coresi, Dimitrie Liubavici, Antim Ivireanul, Dimitrie Cantemir etc. Also there are 123 rare manuscripts from the collections of the Batthyaneum branch in Alba Iulia, from which the most famous is Codex Aureus.

Together with the launching of the new interface of the Manuscriptorium platform, after the process of converting metadata into the new format TEI P5, the National Library of Romania has enriched the digital library Manuscriptorium with another 350 documents from the Batthyaneum branch.

The selection of the documents was made taking into account mainly the age of the book criterion, but also some elements that give uniqueness to some books, such as different artistic bindings or handwritten notes. Also, the state of preservation of the books was taken into account.

For the sixteenth and the seventeenth centuries all the titles from the library collections were digitized, including those printed in Transylvania, while for the eighteenth century only titles listed in the "Romanian Bibliography of Old Books" were selected.

In the historical approach of the Romanian people's culture, the following documents would be representative (Andreescu, 2010):

"The Slavonic Liturgikon/ Liturghier slavonesc", printed at Târgovişte in 1508, by the hieromonk Macarie, is considered to be the first book printed on Romanian territory. The text in Slavonic is printed in two colours: black and red, and the type of letters used is modelled according to the writing of old Romanian manuscripts. Its publication represented the first-time printing of liturgical Mass in Slavonic.

„The Slavonic Apostolos/ Apostol slavonesc", that Dimitrie Liubavici printed in 1547, also at Târgovişte, was a Christian ritual book and it was printed in two editions, one for Valachia and one for Moldavia. Liubavici's printed books are in Slavonic, in black and red ink, and the letters are small and follow each other closely. The few initials that ornamented the book are of Venetian type. The text frontispieces are big, occupying half a page, made of arabesques, in the center of which is the coat of arms of Valachia: the raven with the cross in its beak.

Considered to be one of the most beautiful printed books of Coresi the deacon, the „Gospel for learning/ Evanghelie cu învăţătură", printed in 1581 at Braşov, is richly ornamented with four types of initials, with a full page engraving and new style, humanistic frontispieces. Translated by the priests Iane and Mihai after the Preaching Book from Zabludov, from 1569, the "Gospel for learning/ Evanghelia cu învăţătură" was printed in Romanian, with Slavonic alphabet. Coresi's printings spread in all the Romanian provinces, so the skillful typographer had great merits in the establishment of the literary Romanian language that could be understood in all the Romanian provinces.

Under the guidance of Meletie of Macedonia, in 1640, at Govora, the „Little code of laws/ Pravila mică" was printed, being the foremost accomplishment of this printing press, the first book of laws printed in Romanian, translated from Slavonic by Mihail Moxa. The"Code of laws from Govora/ Pravila de la Govora" was

published twice, once for the Valachia and once for Transylvania.

From the first Moldavian printing press, installed in Iaşi, in the monk cells of the Trei Ierarhi Monastery, the „Romanian book for learning on Sundays all year round and for the Great Feasts/ Carte românească de învăţătură la duminecele de preste an şi la praznice împărăteşti", was published in 1643, translated from Slavonic „in our Romanian language" by the Metropolitan Varlaam, hence the title "The Preaching book of Varlaam/ Cazania lui Varlaam". The book was printed in Romanian with Slavonic alphabet. The printing type is in two colours, with two sets of characters. The ornamental element is very rich, some frontispieces forming religious illustrations in their own right, and the majority have the initials of the engraver.

The printing of the "Key of understanding/ Cheia înţelesului", published in 1678, in the typographer workshops of the Bucharest Metropolitan Seat, is also the work of the Metropolitan Varlaam, who printed in Romanian with Slavonic alphabet the 18 preaching books, translated by the archimandrite Ioanichie Galeatovschi of Cernigov.

A monumental work, the Bible from Bucharest (1688), as it is known, was printed in Romanian with Slavonic alphabet, representing the first complete translation in Romanian of the Old and New Testament. The Valachian voivode Şerban Cantacuzino was the one who initiated the translation in Romanian and the printing of the Bible. This work, whose printing was entrusted to Mitrofan, initiates an intense cultural and typographic activity during the reign of Constantin Brâncoveanu (1688-1714) who, for the reorganization of the printing activity, brings Antim Ivireanu in the country.

In the year 1690 at the Royal printing press in Bucharest the book in Greek was published: „Manual against the Popish Schism / Manual în contra schismei papistaşilor" composed by the

most erudite hieromonk Maxim the Peloponesiac. This most beloved printing of Antim Ivireanu was translated in Romanian and published with Slavonic alphabet at Snagov in the year 1699 with the title „Book or Light/ Carte sau lumină".

In the spirit of the Transylvanian tradition concerning the printing of grammar books, in 1828 "The Romanian Grammar Book/ Grammatică românească" by D. I. Eliade is printed in Sibiu, in Romanian with Slavonic alphabet. In this work, Eliade proposes the simplification of the Slavonic alphabet, reducing it to 28 characters from the initial 33 characters.

In 1795, 107 years after the first printing of the Bible in Romanian, a new edition of the Holy Bible in the national language is printed in Blaj. One of the items in the National Library of Romania collections has numerous manuscript notes from where information regarding the book's price and circulation can be extracted.

The books presented here are available on the portal Manuscriptorium and they are true landmarks of the history of Romanian and Southeast European printing activity, which in the beginning was only in Slavonic, the official language of the Orthodox Church, then in Romanian with Slavonic alphabet; starting from the second half of the 17th century the printing activity diversified, printing books in Greek, and in the beginning of the 18th century even in Arabic.

The methodology of work involved the creation of digital content that is stored on the local server of the library (storage server, accessible via http protocol). Metadata records created by MTool - bibliographic and technical description - offered through the portal Manuscriptorium were uploaded in the central database of Manuscriptorium.

The bibliographic descriptions of the documents are included in Manuscriptorium OCHBR catalog (Open Catalogue of Historical Book Resources). OCHBR groups information

concerning the physical documents in the form of cataloging records in XML format. Manuscriptorium centralizes these records and provides access to the digital documents stored in the databases of its partners.

The description of the digital documents should be as detailed as possible, and its purpose should be to catalogue and create a link between images and an electronic document in the form of a virtual book. (Radu, 2008).

The MASTER standard based on XML allows the creation and use of data and especially semi-structured data, making it adaptable to the diversity of materials to be described and the different search methods. Such record is composed as follows:

1. identification of the document
2. Title
3. contents
4. physical description
5. historical information regarding the document
6. additional information.

MASTER + is an extension of MASTER standard for creating virtual books. Through this a link can be created between descriptive recording and a series of digital images, copies of an original document. To create the mold for a document that has associated digital images the following are required:

1. determination of the number of tabs / page numbering and reporting differences in case there are missing or unnumbered pages;
2. identification of additional pages (if there are unnumbered ones);
3. identification of missing pages

This document can be created using M-Tool program (Annex 2), which allows generation of correct page numbering and links to digital images. Document description is made by filling the following fields and subfields:

- Identification:
  – Shelf number
  – Location - name of the town where the original document can be found
  – Holder - the institution in which the document described can be found
  – Title - title of the document (in the case of miscellanies, the specification of a common title is recommended)
  – Author - the original and complete name of the author
  – Date of publication (copy) - a specific date or a period of time may be mentioned
  – Language - the language in which the document was written (more languages can be specified, if applicable)
  – Notes - any information that is relevant on the document described
- Contents - the field allows a summary of the document's contents
- Illustrations - information on the ornamentation of a manuscript or a printing
- Musical Notation - information on musical notation contained in the document
- Physical Description
  – Link - link description
  – Material - reference material used as support for writing (paper, parchment or a combination)
  – Number of pages - specifying the number of pages or tabs (in this field the missing pages or the numbering errors may be specified)
  – Dimensions
- bibliographic information - reference of publications in which information on the document described have appeared (bibliographies, catalogs, articles, monographs devoted to individual works, etc.)
- Data Printing
  – Place of printing
  – Name of the printer
  – Place of publication
  – Publisher

At the moment, by changing the integrated library software, the bibliographic records can be taken from the library catalog.

Together with the launching of the new interface of the Manuscriptorium platform, after the process of converting metadata into the new format TEI P5, the National Library of Romania has enriched the digital library Manuscriptorium with other documents from the Batthyaneum branch.

The presence of Romanian documents in Manuscriptorium make Romanian cultural heritage accessible and known in the world. Manuscriptorium portal is integrated into the portal of European libraries (The European Library and the European Digital Library), and in this way the documents from the National Library of Romania collections are also present in the latter.

Conclusion

Currently, Manuscriptorium is the largest digital library of manuscripts and old books in Europe providing access to over 5 million pages of manuscripts, old books, maps and other digitized historical documents. The catalog contains over 266 166 bibliographic records of which 41,541 are fully digitized books and the statistics show that they had visitors from 125 countries.

Manuscriptorium partners (www.manuscriptorium.com) by their number, their geographical distribution and the collections held are the evidence of the success of this project. Manuscriptorium Digital Library is integrated into The European Library - TEL through the harvesting process, where data collection is accomplished at the level of bibliographic records (Dublin Core).

The documents found in Manuscriptorium can be also found in Europeana because TEL is the metadata aggregator for the European Digital Library, thus every document from Manuscriptorium and TEL can be found in Europeana.

References

Unesco. Memory of the World. 2012. http://portal.unesco.org/ci/en//ev.php URL_ID=22845&URL_DO=DO_TOPIC&URL_SECTION=-477.html (приступљено 23. јула 2012.)

European Digital Library of Written Cultural Heritage, 2012. http://www.manuscriptorium.com/index.php (приступљено 25. јула 2012.)

Uhlir, Zdenec; Knoll, Adolf. 2009. Manuscriptorium and ENRICH Project: Means for Dealing with Digital Codicology and Paleography. Kodikologie und Paläographie im digitalen Zeitalter = Codicology and Palaeography in the Digital Age. BoD, Norderstedt: 65 -77.

Tîrziman, Elena. 2008. The National Library of Romania contribution to representative European projects – TEL, Manuscriptorium, ENRICH. Les Annales de l'Universite "Valahia" de Târgovişte, Section Sciences Economiques 13(21), Seminaire International Espaces Europeens «Converger dans le cadre de l'Union europeenne» (octobre): 384-392.

Andreescu, Anca. 2010. Colecţiile Speciale. Revista română de biblioteconomie şi ştiinţa informării 6(2): 65-72

Radu, Mariana. 2008. Contribuţia Bibliotecii Naţionale a României la proiectul ENRICH. Manuscriptorium." Informare. Documentare în Biblioteca Naţională a României 1: 83

Dumitran, Adriana. 2009. Proiecte culturale europene: REDISCOVER (2009-2010). Revista Bibliotecii Naţionale a României 15(1-2): 78-79.

Manuscriptorium partners, 2012, http://manuscriptorium.com/index.php?q=content/partners (приступљено 25. јула 2012.)