Jelena Đurović, Milorad Milović
National Library of Montenegro "Đurđe Crnojević", Cetinje
Nenad Jeremić
Blink N Agency, Belgrade

DIGITAL MONTENEGRIN BIBLIOGRAPHY 1494-1994

Summary. In 1979 The National Library of Montenegro "Đurđe Crnojević" launched the Montenegrin Bibliography 1494-1994, a project aiming to compile the most comprehensive list of the overall publishing and printing production of Montenegro, including all bibliographic information on Montenegro and titles written by authors of the Montenegrin origin. This retrospective national bibliography spans half a millennium: from *Octoechos the First Voice*, the first Cyrillic book in Southeast Europe printed in Cetinje in 1494, up to the year 1994. The Montenegrin Bibliography is published in four volumes containing entries for monographs, serials, articles and contributions in serials, and so called 'exteriorica'. So far, 32 books have been published on 15,030 pages of encyclopedic format. They list a total of 229,611 bibliographic entries classified according to the Universal Decimal Classification, provided with appropriate indexes. In 2012 the National Library of Montenegro "Đurđe Crnojević" completed the project entitled The Digital Montenegrin Bibliography. All previously published issues of the Montenegrin Bibliography 1494-1994 have been digitized and made available both online and on DVD, facilitating the use and providing easy access to the enormous amount of bibliographic information. The database has been created by means of optical recognition of scanned images. Its structure is able to support storage of a large number of indexes of various types, and to link them to the scanned material. The database is searchable by different index fields. The search results lead to the corresponding bibliographic record viewed on the scanned page, allowing users to have the complete information displayed within a single window, which makes searching and viewing fast and easy. To meet the needs of reviewing, this system offers appropriate tools for inserting entries and corrections, designed in a way that provides simple, uniform and straightforward usage. PHP, MySQL, Javascript and AJAX are the languages and techniques used in the project.

Keywords: Montenegrin bibliography, history, criteria, structure, digitization.

The Bibliography can be accessed at:
http://www.nbcg-digitalnabibliografija.me/bibliografija_cetinje/

1. The Digital Montenegrin Bibliography Project

The National Library of Montenegro "Đurđe Crnojević", playing the leading role in the development of national library services, launched the project entitled *The Digital Montenegrin Bibliography* in 2010, to complete it in 2012. All previously published issues of the *Montenegrin Bibliography 1494-1994* have been digitized. This project facilitates the use and provides easy access to the enormous amount of bibliographic information. The implementation of the project, which represents a significant step towards a further development of the library information system, as well as a support to the system of scientific information in Montenegro, was funded by the Ministry of Culture.

2. The History of the Montenegrin Bibliography

Retrospective national bibliographies fulfill aspirations to collect and present the overall creative achievements of a nation, expressed in written or another form. They provide a unique source of information for the research of every aspect of social life.

Taking into account that many nations have compiled their national – retrospective and current bibliographies long time ago, considering them to be an essential scientific and cultural need of each developed society, in 1979 the National Library of Montenegro launched the Retrospective Montenegrin Bibliography project with the aim to accumulate in one place – between the covers of a single set of volumes – all the material published in Montenegro and about its people, either in the form of monographs, serials, discussions, articles and literary reviews in serials, or collections of works in both the local and foreign languages(see [1]).

The Montenegrin Bibliography covers half a millennium long period of time, which starts with the first printed Cyrillic book in this part of Europe *Oktoih Prvoglasnik* (*Octoechos the First Voice*) printed in the Crnojević printing press in 1494, and ends in 1994. Apart from the books listed in the first volume, this edition comprises a list of serial titles given in the second volume, covering the period from the first Montenegrin serial *Grlica* issued in 1835 until 1994. The third volume is the most extensive one as it compiles selective lists of articles published in serials. Fifteen books covering the period 1494-1985 have been issued to date, whereas the remaining three typescripts are to be issued soon to cover the period until 1994. The fourth volume contains bibliographies in foreign languages: Russian, Czech, Italian and English. *The Bibliography of books on Montenegro in the languages of nations and minorities from the former SFRY – until 1994* has been prepared and is expected to appear by the end of 2012.

3. Montenegrin Bibliography selection criteria

The Montenegrin Bibliography 1494-1994 project is founded upon a wide scientific basis. Three selection criteria have been used for the inclusion of material: **territorial, national and subject**.

Under the territorial criterion, The Montenegrin Bibliography includes:

- all books, brochures and printed music published on the territory of Montenegro since 1494 regardless of the language, alphabet and author's ethnic origin (exempt from this rule are titles printed for the publishers registered elsewhere – after 1945);
- all books, brochures and printed music published by Montenegrin publishers but printed out of the country;
- all official publications printed in Montenegro, as well as those published by Montenegrin publishers but printed out of the country;
- all serial titles: gazettes, newspapers, journals, almanacs, yearbooks, bulletins and chronicles with supplements, commemorative editions of academies and associations, reports issued over a long period of time bearing continuous numbering, printed or published in Montenegro;
- the selection of articles and contributions in serials which meet the criteria defined for monographs and serials.

Under the national criterion, the Montenegrin Bibliography comprises:

- all works written by Montenegrins and other authors originating from Montenegro published out of Montenegro, regardless of the language and alphabet;
- the works produced by printers of Montenegrin origin printed abroad.

Following the subject criterion, the Montenegrin Bibliography comprises entries relating to Montenegro and its people (more details in [2], [3]).

4. Structure of the Montenegrin Bibliography

The Montenegrin Bibliography is published in four volumes, each containing a number of books / issues depending on the amount of bibliographic material. The volumes are structured in this way:

Vol. I – monographs (book 1, 2, 3, 4, 5, 6, 7, 8, 9, 10);
Vol. II – serials (book 1, 2)
Vol. III – discussions, articles and literary contributions in serials (selective inclusion of works; book 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15);
Vol. IV – 'exteriorica' – works by foreign authors relating to Montenegro and its people (selective inclusion; book 1, 2, 3, 4, 5).

5. Standards and arrangement of bibliographic records, with corresponding aids

Monographs are catalogued according to the international ISBD(M) standard, for serials ISBD(S) standard is used, whereas the entries in the third volume (articles) are catalogued in compliance with the then valid rules applied for the analytical cataloguing of serials. All the material (except Vol. IV) is catalogued and classified according to the UDC system. International standards for bibliographic description are not applied to IV Volume (foreign authors on Montenegro), however, the editors created bibliographic entries according the rules applied in their respective countries. Each book includes indexes of authors, translators, editors, subjects, identified initials, codes and pseudonyms, and lists of references and resources.

Over the past three decades of hard work on *the Montenegrin Bibliography 1494-1994* project, outstanding results have been achieved. Within the four volumes 32 books have been published on 15,030 pages of encyclopedic format. They list a total of 229,611 bibliographic entries which refer to incunabula, paleotypes, books, serials, and contributions in serials. This database represents the most comprehensive treasury of the Montenegrin creativity and an indispensable tool for scientific and other research work.

6. Technological aspects of developed system

Programming tools in digitization of Montenegrin bibliography 1494–1994. The whole process of digitization was organized into following phases: scanning and OCR, parsing of text, making of database, developing of tools for search and view, developing of tools for editing.

In each phase a set of programming tools was developed in one or several programming languages. Most of the codes were written in PHP script language (described in [4]) with JavaScript parts of code to support AJAX functionality (see [5]). Parsing tools for preparing text import into the database were written in Perl script language.

Some of the characteristic tools will be discussed in the text below.

Preparing of text for import into database. One of the hardest and in the same time the most important parts of the project were parsing of text which was made by OCR program. For this purpose, many small Perl scripts were written.

To enable a better search and view performance, it was necessary to recognize different parts of indexes and import them into the database. The text of indexes was created as a result of OCR process applied on scans. As most OCR results, it contains errors. This text is good enough for reading, but it is difficult to automatically recognize separate text fields and import them into the database fields. Most indexes were in Cyrillic and there OCR program made more errors. Another problem was in different format of indexes and the way that entries were marked.

For the purpose of avoiding OCR and format problems, many small Perl scripts were developed to prepare the text before import. Another part of small scripts were used just to import the prepared text into the database.

Indexing of scans. The majority of scans contain bibliographic entries and the main task was to link indexes to entries. The number and length of entries vary on different pages. Entries were indexed by ascending numbers.

For this purpose a PHP indexing script was written. This script works in interaction with the user. The script makes suggestion of indexing for current scans and the user can approve it or correct if needed. Using this script the process of indexing of scans was made five times shorter than manual indexing.

PHP crop tool. Although indexes were turned into text by OCR, for another part of pages, it was needed to optimize them for reading as scans. It included optimizing margin space, which enables a better scan view. The problem was that the scans had very different margin spaces and there was no possibility to make one crop format which could be used for all scans. Also some graphic program for cropping was not an option because it would require too much manual work to achieve cropping of all 15,000 scans.

For the purpose of cropping with different crop offsets, a PHP script was written. It combined cropping functionality with scan preview and dialog where users can set crop offset using numbers rather than mouse marking. Also, it applies changes automatically on groups of scans.

Redaction of indexes. Indexes vary in format and type of fields. There are 14 types of indexes. The main task for the redaction scripts was to make a solution which works with small parts of indexes and enable editing of these small parts. Also, the way of redaction is similar for every index what makes learning of redaction easier.

This functionality was achieved using Javascript and PHP parts of code combined in AJAX approach.

Although these tools have been made for the purpose of this project, some of them can be used in any digitization project.

7. Comparison with another digitization approaches

System for redaction for Serbian bibliography(described in [6]) is suitable for work with more redactors at the same time. Here a person who works on some records can lock it so that

another person who also works on redaction can view it but without the possibility of changing. Many records can be currently locked, but only one record can be manually edited at same time and automatically editing is possible only using import files. In the Montenegro bibliography redaction system, many records can be manually edited at the same time. It is very useful for needs of corrections for mistakes made by OCR. Also, the database is very similar structure as database in Serbian bibliography (as in [6]) . Both databases are MySQL type.

Although complex open source applications as DAM software could be used in digitization (as in [9]), here specific content management system is developed suitable for needs of bibliographies indexes.

Some digitization projects (described in [7]) use well-structured metadata format as Dublin Core. Bibliography project developed own metadata structure according to the needs of indexes.

Literature

[1] Luketic M, MONTENEGRIN BIBLIOGRAPHY, Bibliographic herald 3, pp. 5-41, 1979.

[2] Martinović D, THE MONTENEGRIN BIBLIOGRAPHY, Bibliographic herald 1, pp. 81-93, 1993.

[3] Milovic M, RETROSPECTIVE BIBLIOGRAPHY OF MONTENEGRO : MONTENEGRIN BIBLIOGRAPHY : 1494-1994, Bibliographic herald 1-3, pp. 299-310, 2006.

[4] http://www.php.net/

[5] http://www.w3schools.com/ajax/

[6] Bogdanović M, Jeremić N, *DATABASE OF THE SERBIAN RETROSPECTIVE BIBLIOGRAPHY OF BOOKS: 1868–1944*, Review of the National Center for Digitization 17, pp. 45-49, 2010.

[7] Mijajlović Ž., Pejović N., Ognjanović Z., DOCTORAL DISSERTATIONS IN NCD VIRTUAL LIBRARY, Re-view of the National Center for Digitization 18, pp. 15-18, 2011.

[8] Radovanlija Mileusnić S, *A PROJECT FOR THE DIGITISATION OF A MANUSCRIPT BIBLIOGRAPHIC UNIT – OPPORTUNITIES FOR AND IMPLEMENTATION OF THE VIRTUAL CONTEXTUALISATION OF TEXT*, Review of the National Center for Digitization 21, pp. 127-133, 2012.

[9] Filipović V, Zagarčanin M, Tošić D, Stanišić S, *DIGITALIZATION IN THE BAR COUNTY MUSEUM – PILOT PROJECT*, Review of the National Center for Digitization 23, pp. 57-66, 2013.