

# Historical Newspapers Content as a Base for Scientific Research

## -Europeana Newspapers Project-

Nataša Dakić, Aleksandra Trtovac,

University Library "Svetozar Marković", Belgrade

# University Library "Svetozar Marković"

- University Library "Svetozar Marković" originates from the library of the Lyceum of the Principality of Serbia. Lyceum was founded in 1838 and had a library used by professors and pupils of the Lyceum and students of the Belgrade Gymnasium.
- Based on the decision of the Faculty of Philosophy, the University Library in Belgrade was established in January 1921 and was temporarily given premises at the Faculty of Philosophy.
- Our library building, the first building purposefully built as a library in Serbia, was officially opened on St. Cyril and Methodius Day, 24 May 1926.
- When moving into a newly built building, the collection contained slightly more than 57,000 volumes and the University had 6,000 students.

# University Library "Svetozar Marković"

- Today the library has more than 1.5 million units, and the University has about 90,000 students.
- University Library "Svetozar Marković" is a central library of the University of Belgrade and the parent library for all scientific and academic libraries in Serbia.
- It is one of the founders of the Virtual Library of Serbia, shared cataloguing system which enables access to information from more than 160 Serbian libraries
- KoBSON, a consortium which provides access to 35,000 electronic journals and 40,000 electronic books in full text.



# Digitization process in the University Library "Svetozar Marković"

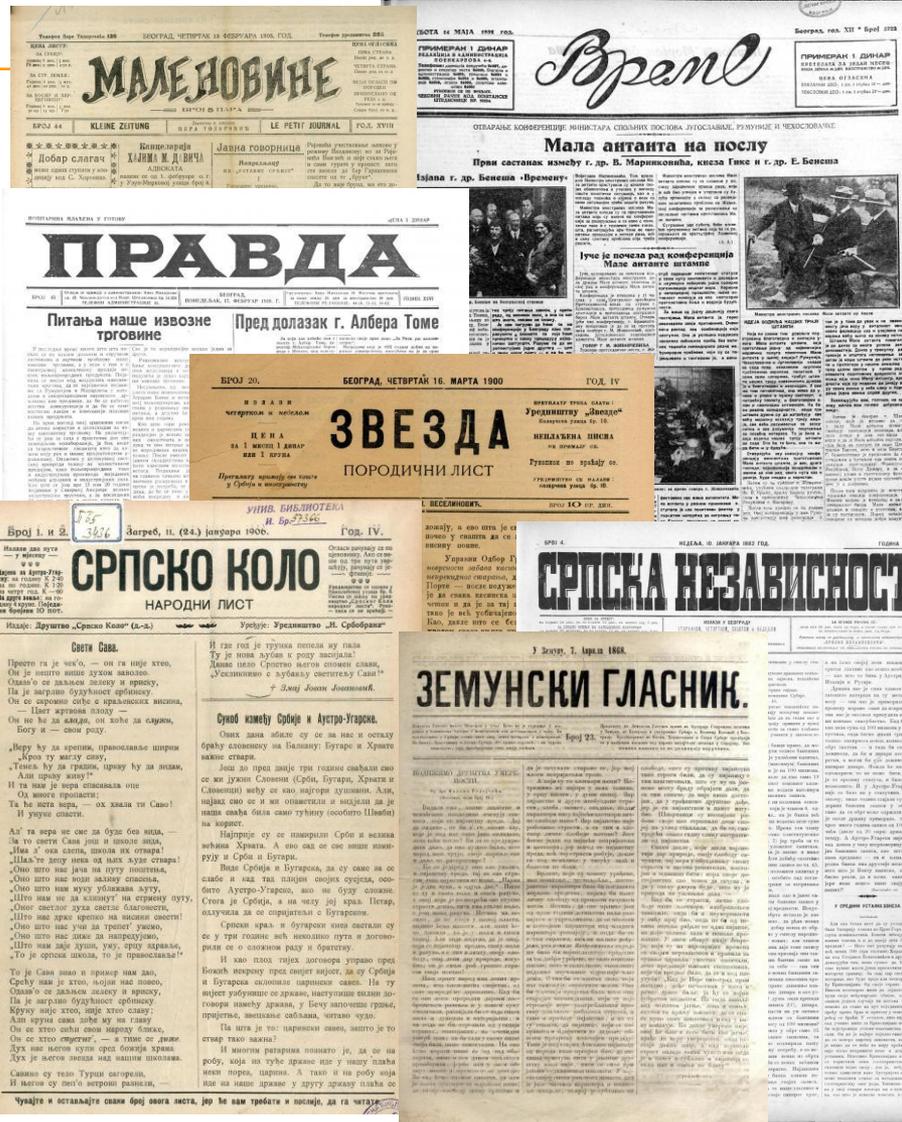
- The process of digitizing old publications began in 2008 with the aim of preservation the old and very valuable collections.
- Digitization has started with cyrillic manuscripts which are located in the Registry of the National Cultural Heritage.
- At first, 129 cyrillic manuscripts dating from the 13th to the 18th century, the oldest ones date from the 13th century had been digitized .
- Digitization continued with old and rare Serbian printed books. The first printed Cyrillic books from our region originates from the 15th century and the oldest one is "Oktoih" printed by Đurđe Crnojević in Cetinje in 1494.

# Digital Newspaper Collection of the University Library

- The printed serial collection of the University Library contains over 8,500 titles in 700,000 volumes. About 5,200 titles are foreign serials, mostly in English, German, French and Russian.
- As newspapers and journals have great significance in scientific research in the fields of history, literature, arts, and social sciences in general, in 2010 the University Library "Svetozar Marković" began the process of digitizing its newspaper collection.
- University Library is a leading institution in the digitization of newspaper collections in Serbia with 47 digitized titles.

# University Library as a project partner in Europeana Newspapers Project

- 45 newspaper titles;
- 392.323 digitized pages;
- issued in the period from 1830 to 1944;
- mostly printed in Cyrillic (one title in Latin), including the old Cyrillic alphabet, which contains Russian letters;
- they will be in open access at the end of the project. All of the newspaper issues will be over 70 years old, which is in accordance with the copyright law.



# Criteria for the selection of the digitized materials

- Titles that are not digitized yet in National Library of Serbia and Matica Srpska Library and could not be found on the portal Digital National Library of Serbia or Digital Matica Srpska Library.
- Availability of the good digital copies which could provide quality information to the end users.
- Technical limitations (poor microfilm copies, missing editions, incomplete supplements). This criteria implies that incomplete years, newspapers lacking extras or bad copies on microfilm were not included.

# The selection of digitized newspapers for Europeana Newspapers Project

- Newspapers – brought about daily political news, information about the way of life in Serbia and the important events of that time.
- Literary newspapers – important Serbian writers of the 19th and early 20th century made first creative steps in literary journals.
- Political and humorous-satirical newspapers – severe criticism of the community, social and political affairs, the current government, the political party policies were truly reflected.
- Newspapers published during the First World War – this is the special group which will enrich a separate part of the European digital collections, called “Europeana collections 1914-1918”.

# Results of newspapers text refinement

- For the content of newspapers provided by University Library „Svetozar Marković”, Optical Character Recognition was planned but for all 392.323 pages.
- University of Innsbruck refined complete UB material with OCR.
- First results of the OCR testing at University of Innsbruck - over 90% success in character recognition.
- Digitized newspapers are mostly printed in Cyrillic, including the old Cyrillic alphabet, which contains Russian letters.

# Manual OCR corrections

- In cooperation with the project partners and the National Library of Serbia, it was agreed that during 2014, University Library project team would perform manual corrections of OCR for the title „Ratni dnevnik“ which contains official reports about the operations in the WWI.
- However, manual OCR corrections require a corresponding software that makes them possible. This work is on stand by for now, while every software used so far changes the structure of the METS-ALTO file.

# A browser for Serbian historical newspapers

- The browser was developed at the University Library "Svetozar Marković" with significant support from colleagues from the National Library of Luxembourg, who shared a tool which they created themselves.
- Beta version of the browser enables a search for keywords within a specified number of newspapers.
- Improvements will require consultation with experts in IT, Serbian language and computer linguistics from the University of Belgrade.

# Browser for Serbian historical newspapers

## Searchable digitized historic newspapers

Select Newspaper

Бодљикаво прасе

Select Year

1942

Select Date

07.12.

Open



University Library "Svetozar Markovic" is presenting the public with the first fully searchable corpus of digitized historic newspaper texts with more than 400,000 pages in Serbian Cyrillic from the period between the turn of the 19th century and the first half of the 20th century. The collection consists of 45 different newspaper titles the most popular of which are *Beogradske opstinske novine*, *Nova iskra*, *Glas naroda*, *Zora*, *Srbadija* and many other.

In order to browse through the digitised pages, it is necessary to choose a newspaper, date and edition.

All newspapers in this collection were digitised within the project *Europeana Newspapers*. University Library "Svetozar Markovic" is one of 18 partner institutions that participate in the project whose aim is to aggregate historic newspaper pages for *Europeana* and *The European Library*. The project especially focuses on the newspapers published during the First World War and in that way in contributes to the project *Europeana Collections 1914-1918*.

Вираница Impressum

YouTube Facebook LinkedIn Twitter

University of Belgrade  
University Library „Svetozar Marković”





Време 1940-08-14



Претрага

Поређајте

Време 1940-08-14

- [-] страна 01
  - [ ] чланак 01
  - [ ] чланак 02
  - [ ] чланак 03
  - [ ] чланак 04
  - [ ] чланак 05
  - [ ] чланак 06
  - [ ] чланак 07
  - [ ] чланак 08
  - [ ] чланак 09
  - [ ] чланак 10
  - [ ] чланак 11
  - [ ] чланак 12
  - [ ] чланак 13
  - [ ] чланак 14
  - [ ] чланак 15
  - [ ] чланак 16
  - [ ] чланак 17
  - [ ] чланак 18
  - [ ] чланак 19
  - [ ] чланак 20
  - [ ] чланак 21
  - [ ] чланак 22
  - [ ] чланак 23
  - [ ] чланак 24
  - [ ] чланак 25
  - [ ] чланак 26
- [+] страна 02
- [+] страна 03
- [+] страна 04
- [+] страна 05
- [+] страна 06
- [+] страна 07
- [+] страна 08
- [+] страна 09
- [+] страна 10
- [+] страна 11
- [+] страна 12
- [+] страна 13
- [+] страна 14

Језик

Лиценца

[Више информација о лиценци](#)

Преглед

# ...ИЗВЕШТАЈИ

## СУТРА:

Јака колебљивост ваздушног притиска. Код нас ће владати променљиво време, претежно облачно са пролазним пљусковима кише и олујама. Температура колебљива.

Текст

...извештадо  
СУТРАг  
Јака колебљивост ваздушног притиска. Код нас ће владати променљиво време, претежно облачно са пролазним пљусковима кише и олујама. Температура колебљива.

# Promotion of content and project objectives

- In order to promote the Europeana Newspapers Project and the achieved results, it was planned that partner institutions, during the project, hold three workshops.
- The first workshop was aimed at presenting the tools to assess the quality and refinement of the text and it was held in Belgrade, at the University Library "Svetozar Marković" on June 13 and 14 2013.
- In addition to the project partners, the workshop was attended by researchers from the University of Belgrade who were presented with the advantages of the digitized newspaper content and the ability to use them in their research work. There were also the representatives of the University of Niš and the National Library of Serbia, who became networking partners in the project after completion of the workshop.

# Marketing activities and media coverage

- All project information are placed on the University Library website, which has an annual visit of over 100,000 users.
- Our library publishes relevant project information on Facebook (7.000 followers) and Twitter (over 150 followers), Newsletter (General Public) - about 2.000 subscribers, Newsletter (Serbian Academic Library Association) - about 400 subscribers, Newsletter (Serbian Library Association) - about 500 subscribers, Press List - about 160 contacts.
- Information about the event was published in all major Serbian newspapers (print and online editions) and online news portals. Moreover, the workshop was the subject of several radio broadcasts on the national radio station.



# Participation in conferences

- Dissemination of the results of the Project has been successfully implemented through numerous appearances on local, regional and international conferences where the Project and its results were illuminated from different perspectives.
- Since the beginning of the project in 2012 until today librarians of University Library involved in Europeana Newspapers Project participated in 12 conferences.
- After the paper presentations, we received many questions from the participants who were interested in the possibility of using the digitized materials. These contacts are extremely important particularly when it comes to the colleagues in the region. Due to the common Yugoslav past, Serbian language and Cyrillic are not an obstacle, and the digitized collection of Serbian historical newspapers is highly important for researchers from all over ex-Yugoslavia.

# Production of promotional materials

The project team launched another marketing campaign in order to encourage our fellow librarians and users to start to use digital newspaper collections more actively. Bags, cups, mugs, notebooks and folders featuring the Serbian historical newspapers, either in terms of titles with beautiful old fonts or various lavish illustrations have been designed and manufactured.

- The first part of the promotional materials was distributed to the workshop participants in Belgrade, and the other part is being distributed to the participants of various events at our institution.



# Further activities

- Promoting the content,
  - accurately determining the target user groups and
  - training them to use this valuable content.
- 
- In the third and final year of the project, our goal is to bring as closely as possible the project and digital content to the researchers in Serbia by organizing lectures, workshops and presentations for the members of the scientific community of Serbia, but also for the librarians of academic libraries.

# Conclusion

- Newspapers represent a valuable and an indispensable source for scientific research because the prominent poets and scholars primarily published their literary, artistic and scientific works in newspapers. This content will be available completely free of any legal restrictions at the Europeana portal.
- We hope that researchers and academic librarians in Serbia will recognize the importance of this project and Europeana portal will become a starting point in their research work.





**Thank you for your attention!**

[aleksandra@unilib.bg.ac.rs](mailto:aleksandra@unilib.bg.ac.rs)

[dakic@unilib.bg.ac.rs](mailto:dakic@unilib.bg.ac.rs)