

**Aleksandar Mihajlović, Vladisav Jelisavčić,  
Bojan Marinković, Zoran Ognjanović**

Mathematical Institute  
of the Serbian Academy of Sciences and Arts  
Belgrade, Serbia

**Veljko Milutinović**  
School of Electrical Engineering  
University of Belgrade, Serbia

## **AN OVERVIEW OF AND INNOVATIVE PERSPECTIVES FOR THE SERBIA-FORUM CULTURAL HERITAGE DIGITIZATION PROJECT**

**Abstract:** The initiation of the Serbia-Forum project, geared towards the digitization, presentation and organization of articles of Serbian cultural heritage and historic significance began successfully in March 2012. Strides made to collect as many significant digitized articles proved to be successful. Digitization of significant media content using state of the art scanners has been performed and is in constant progress. The Serbia-Forum is structured according to the Austria-Forum model. The structure is exhibited via the presentation of digitized content on the web. The Mathematical Institute of the Serbian Academy of Arts and Sciences or MISANU, the instigator of the Serbia-Forum project, in cooperation with many partner institutions both in and out of the Serbian government has enriched its collection of digitized content with 80,000 new units. Currently, the content ranges from postcards, newspapers, photographs and books to other relevant media. Progress has been made in content organization and searching through the content structure. Metadata enrichment and dynamic or interactive and user defined metadata development methods for richer content description are under development. The enriched metadata will aid in the semantic search capabilities or services that the Serbia-Forum is to provide to its users. This will enable easy, fast, flexible and user friendly content searching experience. Developments in this direction are planned for 2014.

**Keywords:** Serbia-Forum, Austria-Forum, Java Server Pages Wiki, National Center for Digitization metadata standard

### **1. Introduction**

The beginning of the twenty first century has seen a rapid increase in the use and availability of computing technology. In addition to this, the continuous growth of the Internet's availability gave birth to a number of new and revolutionary data sharing ideas.

Among these ideas is the “knowledge across the wire”, whose goal initially was to reduce frequent trips to libraries, book stores, and newspaper stands and give people quick and easy access to the information they need, when they need it [3]. Initially, prior to the emergence of the Internet, standalone installations were the forerunners of data storage. The high purchase and maintenance costs of legacy standalone systems initially did not aid in stressing the importance of digital storage of cultural heritage data and other knowledge, i.e., academically relevant information such as voluminous encyclopedias. The emergence of the Internet made storing these volumes of information a much simpler task. With Internet, one can distribute the knowledge across many interconnected and inexpensive computers. Internet facilitated the presentation, storage and preservation of information at much lower hardware and maintenance costs. One such web application, under constant development since the year

2012, geared towards the presentation, storage and preservation of culturally relevant data is Serbia-Forum [3].

## 2. Heritage preservation

To recapitulate, some major factors influenced the facilitation of inexpensive and easy knowledge exchange:

- Inexpensive high volume computer manufacturing technologies
- Advances in integrated circuit design which led to
- Smaller computers, large volumes of inexpensive storage, i.e., storage media, Internet and the Web.

For the sake of clarity, the Web is an application that is present on the Internet. It can be said that it presents the user with a logical view or an abstraction of physically networked computing devices. It is presented as a single application entity via the web-client application, more commonly known as the web-browser. The web-browser is the gateway to the experience called web surfing. The major factors of knowledge exchange facilitation had a profound impact on the content commonly found in museums and archives.

National heritage in its physical form is located in “safe houses” such as archives or museums. Physical preservation of national heritage is an unrewarding task provided the fact that all physical objects have their own period of decay. Additionally, national heritage may be exposed to threats of natural disasters, fires, wars and other forms of human manipulation or error. These potentially harmful events for heritage artifacts stimulate the need to digitally preserve and store copies of heritage, i.e. culturally and historically significant content found in the mentioned institutions.

## 3. Characteristics of the Serbia-Forum

The Serbia-Forum is a web application with the goal of preserving digitized units of cultural heritage which are significant to the Republic of Serbia [3]. The web application is based on dynamic content generation and presentation delivered by the JSP Wiki framework, running on the Tomcat 7 web server. From the beginning of its development in March of 2012, the web application was built to serve two specific purposes. Firstly, it serves as an encyclopedic collage of articles written by credible authors. The credibility of authors is checked, along with the sources of information within the articles. If an article is duplicated, that is if it is written twice by two different authors, then both articles are saved for the sake of recording article progress. This allows users, i.e. readers of these articles to track certain changes and to observe different points of view of authors. Secondly, the application serves as a hub or centralized node for collecting quality controlled, digitized units of cultural heritage content from other distributed collections, i.e. archives and museums.

The Serbia-Forum web application is based on the existing forum web application, the Austria-Forum, developed by Prof. Dr. Hermann Maurer at the Technical University in Graz, Austria. The portal based functionality of the application facilitates navigating through content and the contents respective presentation. The Serbia-Forum works according to two sets of four axioms, termed primary and secondary. The secondary axioms are those that make the Serbia-Forum part of the forum family promoted by the Austria-Forum model. The secondary axioms are:

- (1) Content within the portal is semantically searchable via adequate metadata.
- (2) The digitized content comes from credible sources (government controlled institutions) and other article based content is accompanied by authors' biographies.
- (3) If the content is changed, every version is tracked and can be viewed by the user.

(4) Integrity of the content is protected.

All forums abide by the secondary set of axioms presented by the Serbia-Forum. The primary axioms are unique to the Serbia-Forum and are precisely what makes it different from all other forums in the works. The primary axioms are:

- (1) National heritage is owned by national institutions, and is exposed by servers under control of these institutions!
- (2) National heritage falls under a number of different legal regulations, not only Creative Commons!
- (3) Stress is made on quality and ranking, not quantity and chaos!
- (4) Possible semantic culture-oriented translation into a number of relevant foreign languages!

The secondary set of axioms is global in nature and not exclusively linked to the Serbia-Forum [3]. This set of axioms makes the Serbia-Forum a Forum based application on the standards presented in Austria-Forum. The primary set of axioms is local in nature and bound to the Serbia-Forum project [3]. Current forums in development under the secondary axioms are Austria-Forum, Serbia-Forum, German-Forum, English, Dutch and other forums. According to these germinating forums we can conclude that each nation has a need to preserve its cultural heritage digitally.

#### 4. Progress during 2013

**4.1. Equipment.** The year 2013 was very progressive for the Serbia-Forum project. The team and the project development leader of the Serbia-Forum are in the Mathematical Institute of the Serbian Academy of Sciences and Arts or MISANU, located in Belgrade, Serbia. Since the onset of the Serbia-Forum project in early March of 2012, MISANU obtained several units of equipment necessary for digitization and content storage and presentation. The four main units obtained were two different scanners for digitization of print resources, a racked server and storage media or backup media. One of the scanners is a stationary ATIZ scanner with a “V” shaped book frame or book platform. The benefits that this particular platform brings are geared towards older books which may have damaged spines [1]. Opening a book at an angle of 180 degrees flat on a tabular surface may heavily damage it and render it unusable or permanently disfigured. This is something that needs to be paid attention to. The scanner utilizes two digital photographic cameras that take pictures of each page of an opened book on the V frame. Physical size can range up to the A2 paper format, i.e. size. The images are rendered, digitally modified and stored on the hard drive. Usually, these images are stored in the form of image RAW files, which are more or less generic high quality image files used by Cannon photographic cameras. This offers the digitization team the flexibility to change from RAW file format to any other image file format or even PDF format. The other scanner is a customized portable VESTIGIA scanner developed at the Technical University in Graz. This scanner utilizes a similar V shaped frame. However, it does not utilize two cameras, rather only one camera. This doubles the time required to scan a book in contrast to using the stationary ATIZ scanner. The VESTIGIA scanner is useful for scanning content or books that are not mobile or are in too poor condition to be transported to MISANU for scanning on the stationary ATIZ scanner. Such books include old medieval manuscripts found in monasteries located considerably far away from MISANU or Belgrade. From the content presentation aspect, a racked server is used with a growing number of terabytes of space running on a Linux operating system and utilizing an Apache Tomcat 7 web server. This is the host machine for the Serbia-Forum project. Several backup systems are present throughout the

local geographic region to ensure constant presence on the Web. Finally, a RAID based storage module was obtained, EMC2 VNX5300, to store, backup and keep record of the main copy of the Serbia-Forum web-application locally.

**4.2. Application progress.** In addition to the equipment obtained, the Serbia-Forum team managed to externalize the Serbia-Forum application from its template application – Austria-Forum. This means that most of the dynamic functionality modules were translated from the original German into Serbian. Additionally, preliminary steps have been taken to develop a firm basis for semantic content search. These steps include the implementation of a detailed metadata standard known as the NCD Metadata standard [4, 5] named after the National Center of Digitization in Belgrade, Serbia. The implementation of this standard will help facilitate correct or at least relevant query hits for semantic search algorithms under current development. A new module for presenting and creating metadata has been implemented. Due to the specific nature of digitized cultural heritage content, that is the:

- 1) Vast diversity of digitized content types,
- 2) Need for preserving complex content structure and
- 3) Constantly increasing pool of knowledge,

Special attention was given to enable dynamic addition of metadata for new forms of digital heritage as well as structuring and updating metadata for the existing ones. In order to do so, a XML database solution was chosen and both administration and presentation layers were generated automatically. New forms of digital heritage, thus, can be easily added to a database in the form of a XSD schema, and appropriate views would be automatically generated. These views include management forms (user interface for adding metadata for new digital objects as well as editing metadata of the existing objects), metadata presentation view (various presentations of metadata to the ordinary user) and metadata search forms (for making metadata queries). All three views are generated based on the content schema.

Forum software has been adapted in order to allow multiple localization views of every digital item. Every cultural heritage item can have its digital representation in any language (currently in Serbian and English only, but new languages can be dynamically added). A new editor interface is planned in order to allow users comfortable translation of digital presentation pages between languages. Users are active contributors to the Serbia-Forum. They contribute digitized content and/or their original written articles. Forum software enables users to create their own context for every digital heritage item they want to present. This is already accomplished by using wiki-style markup and numerous plugins that are available for users and content providers (e.g. National Library of Serbia, Serbian National Archives) as a tool to present their content. Users can write their own presentation pages for their digital objects. Currently a broad toolset is already at users' disposal, and many new plugins are on the way. New tools for facilitating the editing and manipulation of presentation pages were developed for the users. Un-experienced users can now use a graphical interface when presenting their content, without having to learn wiki markup. The presentation of audio heritage files has been introduced to the Forum. A new tool for adding audio digital content was created.

Strong collaboration with local institutions in Serbia is being enforced. One institution in particular has proved to be pivotal player in the fast development of the Serbia-Forum, the National Library of Serbia (NBS). Since collaboration was established between NBS and MISANU, the Serbia-Forum received its first major content contribution of approximately 1,800,000 pages of digitized material. Currently, another large content contributor is the Museum of Applied Arts in Belgrade, Serbia. It has already contributed whole collections of

digitized art, church/monastery frescoes from the 12<sup>th</sup>, 14<sup>th</sup> and 15<sup>th</sup> centuries and many other religious and secular works of art. The purpose of their contribution is to digitize their exhibitions and promote the museum by utilizing the presence of the Serbia-Forum.

**4.3. Digitization progress.** The second most significant partner of MISANU next to NBS is the Archives of Serbia (AS), which is still in the early stages of digitization. Since the first days of collaboration MISANU actively took part in scanning several books and archival documents using the portable scanner VESTIGIA in AS, among them being the 1860 census in Serbia, which appears in over ten volumes of names, signatures and other personal data. In its own digitization initiative MISANU has successfully digitized almost all of the volumes of the famous local journal of mathematics, YUJOR.

## 5. Conclusion

The aim of the Serbia-Forum is to unify a community of credible authors to continually write articles for the Forum and to provide high-quality, trustworthy cultural heritage content as well as to serve as the central node for collecting quality units of digitized content of cultural heritage significance [3]. In its first year of active service the Serbia-Forum has made several strides towards achieving its goals. Strides have been made in the areas of logistics, application development and sole digitization of content as well. New partners have been made and several significant contributions have aided in the Serbia-Forum's fast growth, in terms of content richness. A new metadata standard has been implemented which will prove to be pivotal for future semantic search algorithms. For a year's worth of operation, development of the Serbia-Forum has reached all of its expectations and the goals are being realized according to an active schedule.

## References

- [1] ATIZ *BookDrive pro. The most powerful book scanner for large digitization.* ATIZ Innovation Co., 2013.
- [2] Maurer, H., Milutinović, V., Ognjanović, Z. *The Serbia Forum*, ppt presentation MISANU 2012.
- [3] Mihajlović, A., Jelisavčić, V., Marinković, B., et. al. *The Serbia-forum Cultural Heritage Digitization Project with Emphasis on Semantic Indexing.* Review of NCD, vol. 22, 2013, 47-54. Available at:  
<http://elib.mi.sanu.ac.rs/files/journals/ncd/22/ncd22047.pdf>
- [4] Ognjanović, Z., Butigan-Vučaj, T., Marinković B. *NCD Recommendation for the National Standard for Describing Digitised Heritage in Serbia.* Metadata and Semantics, Springer, 2009, 45-54.
- [5] Ognjanović, Z., Butigan-Vučaj, T., Marinković, B. *Predlog Nacionalnog standarda opisa pokretnih kulturnih dobara.* Review of NCD, vol. 11, 2007, 1-11. Available at:  
<http://elib.mi.sanu.ac.rs/files/journals/ncd/11/ncd11001.pdf>

[mihajlovic@mi.sanu.ac.rs](mailto:mihajlovic@mi.sanu.ac.rs)  
[vladisavj@gmail.com](mailto:vladisavj@gmail.com)  
[bojanm@mi.sanu.ac.rs](mailto:bojanm@mi.sanu.ac.rs)  
[zorano@mi.sanu.ac.rs](mailto:zorano@mi.sanu.ac.rs)  
[vm@etf.rs](mailto:vm@etf.rs)