Serbia Forum –
Digital Cultural Heritage Portal

Aleksandar Mihajlović[1,*], Vladisav Jelisavčić[1], Bojan Marinković[1], Milan Todorović[1],
Zoran Ognjanović[1], Siniša Tomović[1],
Vladimir Stojanović[1], and Veljko Milutinović[2]

[1] Mathematical Institute of the Serbian Academy of Sciences and Art
Knez Mihailova 36, 11001 Beograd, p.p. 367, Serbia
{mihajlovic,vladisavj,bojanm,mtodorovic,zorano}@mi.sanu.ac.rs,
{sinisatom,vladimir.d.stojanovic}@gmail.com
[2] School of Electrical Engineering, University of Belgrade,
Bulevar Kralja Aleksandra 73, 11000 Beograd, Serbia
vm@etf.rs

Abstract. Serbia-Forum is a web application portal designed and implemented by the Mathematical Institute of the Serbian Academy of Sciences and Arts (MISANU) whose goal is to digitally make available, many units of cultural heritage belonging to the national heritage of the republic of Serbia. The Serbia-Forum project as a whole is geared towards the digitization, presentation and organization of digitized Serbian cultural heritage. In the past two years since the onset of the Serbia-Forum project, MISANU 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. Serbia-Forum is growing successfully both in impact and in content. Currently, the portal content ranges from postcards, newspapers, photographs, books and other relevant media.

Keywords: Serbia-Forum, Austria-Forum, NBS, AS, VESTIGIA, ATIZ.

1 Introduction

With computers getting smaller and smaller, i.e., more portable, the number of interacting users and the frequency of their usage increase proportionally. Today one need not use a desktop computer to search the web. One may simply pull out his or her smart phone running a mobile phone based operating system such as "Android" or Apple's "iOS" which is connected to the internet via a 4G ultra-broadband wireless connections and search for content. Vast quantities of information are now available with a few clicks of a mouse or a few taps on a touchscreen phone. The presentation of information and its exchange is facilitated through content presentation web technologies such as XML and XSLT technologies used for document modeling. In addition to the technological advancements of web technologies, the incessant growth of the Internet's availability and affordability the number of users of Internet based services has increased. Every user of Internet is a contributing entity to the growth of the

collective knowledge base that Internet presents. This mind set gave rise to a new and revolutionary data sharing idea, "knowledge across the wire".

Internets ability to preserve information has a profound impact on content commonly found in museums and archives. With the threat of natural disasters, fires, wars and other forms of human unethical manipulation or error, the need to preserve and store digitized copies of tangible units of culturally and historically significant content found in important institutions has become feasible. The facilitation of large scale digitization projects particularly are important for western Balkan countries such as Serbia, provided the country's turbulent track record from the last hundred years (several hundred years). Not only does Internet preserve, but also it facilitates the availability of archival content to an enormous number of users who can search through content from their homes or offices. One such web application, under constant development since the year 2012, geared towards the presentation, storage and preservation of culturally relevant data using Internet is Serbia-Forum [1].

2   The Forum Concept

The Serbia-Forum (*www.serbia-forum.org*) is a web application with the goal of preserving digitized units of cultural heritage which are significant to the republic of Serbia [1, Fig. 1]. 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 is built to serve two specific purposes.

Fig. 1. The Serbia-Forum homepage

Firstly, it is to serve as an encyclopedic collage of articles written by credible authors. The credibility of the authors is checked, along with the sources of information within the articles. If the article is duplicated, that is, if it is written twice by two different authors then both articles are saved and are presented. 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 is to serve as a hub or centralized node for collecting quality controlled, digitized units of cultural heritage content from other distributed collections, i.e. such as archives and museums.

The web application itself is based on the existing forum web application, Austria-Forum (*www.austria-forum.org*), developed by Prof. Dr. Hermann Maurer at the Technical University in Graz, Austria. Given the technological similarities of Serbia-Forum and Austria-Forum mentioned, the principles of these two applications are nonetheless very different; discussed below are the differences between the primary and secondary axioms that govern the mission of Serbia-Forum. Navigation through the digitized content is facilitated by the portal based functionality of the application. Each of the portals of the web application is presented via an application plugin addition programmed using JSPWiki. Portals can be freely added or removed, making the administration and design of the application flexible to the addition of new content types i.e. collections or aesthetic-cosmetic changes. Currently many fori applications are in the works [2]. Serbia-Forum works according to two sets of axioms, termed primary and secondary.

Secondary Axioms:
(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 biographies of the author
(3) As the content is changed, every edit is tracked and can be viewed by the user
(4) Integrity of the content is protected

Primary Axioms:
(1) National heritage is owned by national institutions, and is exposed by servers under control of these institutions!
(2) National heritage falls into 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 are those that make Serbia-Forum part of the fori family promoted by the Austria-forum model. These axioms are global in nature and not exclusively bound to Serbia-Forum [1]. This set of axioms makes Serbia-Forum a Forum based application by standards presented in Austria-Forum. All future forums should follow the secondary set of axioms. The primary axioms on the other hand are specific to Serbia-Forum and are responsible for the uniqueness of Serbia-Forum [1]. Current fori in development under the secondary axioms are Austria-Forum, Serbia-Forum, German Forum, English, Dutch and other fori. According to these germinating fori we can conclude that each nation has a need to preserve its cultural heritage digitally.

3   Equipment

The Serbia-Forum project is comprised of both content digitization and presentation tasks. The digitization equipment and the presentation equipment was selected carefully in order to suite the special needs of the project, which are in essence influenced by diverse working conditions, both mobile and stationary in nature, and computing requirements necessary for fast and dependable digital content retrieval and storage. The digitization tasks are performed using state of the art scanning equipment suited for various regional i.e. working environment conditions. The content presentation tasks are performed on several computing modules i.e., servers, switches etc., wired to a T1 gigabit academic broadband line.

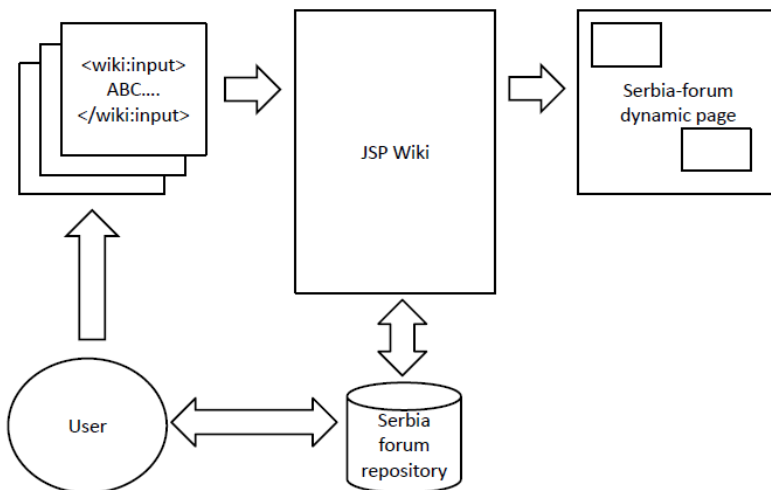## 3.1   Digitization Equipment

The hardware used by the Serbia-Forum project includes both system oriented and scanner oriented devices for presentation and digitization of relevant content. The two main digitization devices at our disposal are two different scanners for digitization of hard copy content. For scanning larger format paper documents, a stationary ATIZ book pro scanner with a "V" shaped book frame is used while for scanning non-mobile content a smaller mobile VESTIGIA scanner is used for scanning documents onsite. The larger ATIZ scanner's "V" shaped book frame retains the integrity of older books during the scanning process, especially those which may have damaged spines [3]. Opening a book at an angle of 180 degrees on a tabular flatbed surface may heavily damage the book and render it unusable or permanently damaged. Physical content size can range up to the A2 paper format. For scanning speed and efficiency the scanning is performed by two high quality Cannon EOS cameras with 300 dpi resolution that can scan two pages of a book at a time. These cameras have proven to be top of line photography devices and are able to reach the scanning quality of conventional flatbed scanners without heavily damaging the paper or content of the document being scanned. Upon scanning the images are digitally rendered and processed either manually or automatically using a plethora of image editing software. All RAW copies of images are saved in case anyone image need be converted into another format. This offers the digitization team the flexibility to change from RAW file format to any other image file format or even PDF format. The VESTIGIA scanner being portable in nature also utilizes a similar "V" shaped frame. However, it doesn't utilize two cameras, rather only one camera. This increases the time required to scan a book in contrast to using the stationary ATIZ scanner by a factor two. The VESTIGIA scanner is useful for scanning content or books that are not mobile or are in very poor condition to be transported to MISANU for scanning on the stationary ATIZ scanner. Examples of such books include old medieval manuscripts found in monasteries that cannot be transported to MISANU for scanning due to their fragility, rareness or everyday religious/ceremonial importance.

3.2   Presentation Equipment

In terms of content presentation hardware, a racked server is used with a growing number of terabytes of space running on a Centos Linux operating system and utilizing an Apache Tomcat 7 web server. The selection of the server itself was influenced by the technology used for developing the application, the JSPWiki engine. The JSPWiki architecture is tailored for the Tomcat 7 server. The decision to use JSPWiki was made in order to standardize similar current and future forum applications. According to preliminary project research, we have discovered that web applications with similar project goals use the JSPWiki engine, such as the Austria-Forum, which has become a crucial partner in the development of Serbia-Forum. This is the host machine for the Serbia-Forum web application. 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   Technical Achievements

The Serbia-Forum JSPWiki engine is the backbone of the heritage portal. It enables the user to dynamically create webpages on the server end. The pages are created by the Java Virtual Machine or JVM on the web server end, creating the illusion of platform neutrality where the user need not have a JVM installed on the client machine in order to render pages. This concept makes the Serbia-Forum heritage portal a dynamic web application which reduces both memory and processor strain on the client computer. The JVM accompanied by an XML eXist database makeup the structural modules of the web server. The eXist database is used for content storage and organization. The database is very flexible in nature and gives a lot of moving room for

Fig. 2. Functionality schema of the Serbia-Forum web application

the application developers in terms of content and metadata categorization, classification and other important organizational tasks necessary for the implementation of semantic content searching mechanisms. The JSPWiki engine enables effective collaboration between the users of the Serbia-Forum portal, allowing each registered member to dynamically create their own content dedicated pages (see Fig. 2). The way the JSPWiki engine is utilized starts first with user registration which is followed by content uploading and categorization. The second step is forming the page layout of the using wiki syntax. This page is the user's presentation on the Serbia-Forum heritage portal which allows the user and his/her/it's content to be adequately represented on the portal. The wiki syntax then instructs the JSPWiki engine on the server side to assemble the user's page according to the syntax and to link any relevant content of the user, located in the repository, with the user's page. Once content linking is complete, the dynamic page is ready for viewing.

The Serbia-Forum team managed to externalize the original JSPWiki application from its template application Austria-Forum. The dynamic functionality modules present in the Austria-Forum were translated from the original German into Serbian. Preliminary steps have also been taken to develop ground for semantic searching of content within the web application. Due to a specific nature of cultural heritage, that is:

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

A new module for presenting and creating metadata has been implemented respectfully. This module abides by a new and detailed metadata standard known as the NCD Metadata standard [4,5] or (National Center of Digitization Standard), named after the National Center of Digitization in Belgrade Serbia. This standard will facilitate correct or at least relevant query hits for semantic search algorithms which are currently under development. Within this new standard special attention was given in order to enable dynamic addition of metadata for new digital heritage types as well as structuring and updating metadata for existing ones. In order to do so, a XML database solution was chosen and both administration and presentation layers are generated automatically. New digital heritage types, thus, can be easily added to a database in 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 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.

The Forum software has been adapted in order to allow multiple localization views of every digital item. Now every cultural heritage item can have its digital representation in any language (current content is 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. The Forum application

enables users to create their own context for every digital heritage item they want to present. This is already accomplished using wiki-style markup and numerous plugins that are available for users and content providers (e.g. National Library of Serbia, Serbian archives ...) as a tool to present their content on the web. Users can write their own presentation pages for their digital objects. Currently a broad toolset is already at disposal, and many new plugins are on the way to ease the user's experience.

Strong cooperation with local institutions in Serbia is being maintained. One institution in particular has proven to be a pivotal contributor in the development of Serbia-Forum, the National Library of Serbia (NBS). Serbia-Forum received its first major shared content contribution of approximately 1,800,000 pages of digitized material from NBS. Currently, another large content contributor is the Museum of Applied Arts in Belgrade, Serbia. It currently has contributed whole collections of digitized art, church/monastery frescoes from the 12th, 14th and 15th centuries and many other religious and secular works of art. This synergistic approach aids in increasing the exposure of cooperative institutions and increasing the quality controlled content base of Serbia-Forum.

5   Conclusion

The aim of 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 a central node for collecting quality units of digitized content of cultural heritage significance [1]. Serbia-Forum hopes to be a research tool enriched with both open (mildly controlled user written articles) and highly quality controlled content. Additionally, Serbia-Forum aspires to become an integral player in the digital preservation process of the republic of Serbia. A new metadata standard has been implemented which will prove to be pivotal for future semantic searching algorithms. For a year's worth of operation, development of Serbia-Forum is reaching all of its expectations and goals are being realized according to an active schedule.

References

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