

Dragan Golubović

Mediacentar,
INFOBIRO digital archive,
Sarajevo

DIGITALIZATION OF BIH NATIONAL MUSEUM'S GLASNIK (HERALD), ITS PRESENTATION AND DIGITAL PRESERVATION

Abstract. Digitization of cultural heritage and access to digitized collections form part of the established EU standards. Printed material, as part of cultural heritage, represents a great challenge in the digitization process. Historically, sociologically and culturally, press content appears as an interesting material. The author presents a model of digital presentation of National Museum's Glasnik through INFOBIRO digital archive and explores the system of digital preservation.

Key words: digitization, INFOBIRO digital archive, National Museum Glasnik, users, digital paper

Introduction

Through its INFOBIRO digital archive, Mediacentar Sarajevo enables a full online search of digital copies of BiH newspapers and other documents. In addition to Glasnik, INFOBIRO also archives digital copies of NADA, Bosanska Vila, Sineast, Naši Dani, and Oslobođenje. This paper aims to explore Mediacentar's experience with digitization, digital presentation and preservation. INFOBIRO has been archiving and preserving printed material for many years. Its staff members have undergone the entire archiving and preservation process – from compiling and binding material, to microfilming and digitizing its content, thus protecting it and making it accessible to a circle of users as wide as possible.

Digitization of National Museum's Glasnik

National Museum's Glasnik (hereinafter Glasnik) is the oldest scientific publication in BiH. Its first issue came out in 1889 and the magazine is still officially in existence. Glasnik treasures vital information on archaeological localities in BiH and the region, as well as valuable historical, ethnological and scientific content. Glasnik is a unique source of information written by experts of the day. It was produced quarterly both in Latin and Cyrillic script, in 25,5x16,5 format.

Process of Glasnik's digitization includes several steps:

1. compiling all issues (1889–1940)
2. migrating content from paper into digital form
3. processing and archiving material in INFOBIRO digital archive at www.infobiro.ba
4. producing DVD

Copyrights. Since Mediacentar originally did not possess Glasnik in its archive, it was necessary to sign an agreement on cooperation with the National Museum and obtain approval

for digitization of Glasnik and its digital publication online. Copyrights are an important aspect of digitization. The agreement on cooperation between the National Museum and Mediacenter confirmed copyrights of the museum and allowed Mediacentar to digitize Glasnik and make it available online through INFOBIRO.

Legibility of digitized content. Legibility of digitized content is the main goal of any digitization process. Each individual page needs to be scanned and its content transferred from paper into a digital format. The digital format used was PDF and scanning was done in 350 DPI resolution.

The digitization of Glasnik included not only the conversion of its content into digital format, but also extraction of word text from PDF file, which made the content most accessible for further processing. This procedure was carried out with the help of OCR technique (*Optical Character Recognition*). After each page was individually processed in PDF Transformer version 2.0, Word text was extracted and prepared for subsequent archiving in INFOBIRO digital archive.

OCR technique is the most reliable way of retrieving word text from a digitized document. PDF Transformer version 2.0 supports both BiH scripts (Latin and Cyrillic). The resulting word document matches with the original in the range between 60% and 70%. It is necessary, however, to review the document and correct any mistakes made by the system. The retrieval of the original text marks the end of one part of digitization process.

Presentation and accessibility of National Museum's Glasnik. Every individual text archived in INFOBIRO has a description attached according to the following parameters: text heading, title, subtitle, publication number, date of publication, page number and key words in the case of interview. Users thus have complete overview of all information related to any individual text at any moment.

Izvadak iz ljetopisa fra Nikole Lašvanina,

po prijepisu fra Stjepana Margetića-Jajčanina (čuva se. u franjevačkom manastiru Gorica kod Lijevna.)

Download: pdf 

Autor: STJEPAN MARGITIĆ-JAJČANIN

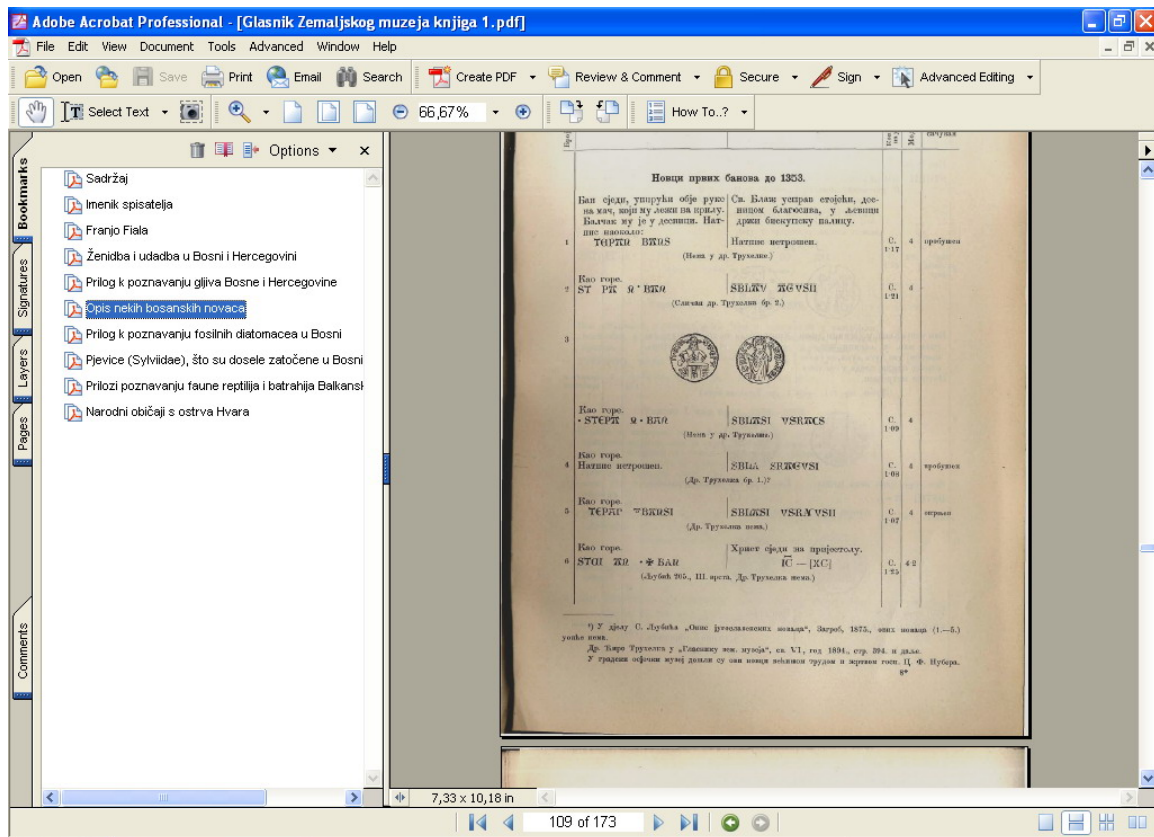
Detalji

Autor(i):	STJEPAN MARGITIĆ-JAJČANIN
Publikacija:	GLASNIK ZEMALJSKOG MUZEJA
Datum izdavanja:	01/07/1889
Tip sadržaja:	MEDIJSKI SADRŽAJ
Broj stranice:	77
Ključne riječi:	[N/A]
INFOBIRO arhivski materijal broj:	397778
Datum dodavanja u arhivu:	07/12/2007
Kredit(a):	Tekst je besplatan

Each text is entirely searchable through key words, which is became possible thanks to word text we previously extracted from PDF file. Both word text and PDF version of the page where the text is located are available. This enables users to search Glasnik and to see its original layout as published over a century ago. Accessibility is an additional bonus – regardless of his or her location in the world, any user can access the publication at any time online through INFOBIRO database.

System of digital preservation. The preservation of digital copies is equally important in the digitization process. Since quality and longevity of media such as DVD and CD cannot be guaranteed, the preservation of digital content in this manner is not the most reliable option. In INFOBIRO, digital copies are preserved in three different ways, both online and offline.

Firstly, each digitized newspaper is copied on DVDs and archived in other libraries and archives in BiH, which gives additional security. In the DVD format of the digitized Glasnik all articles are indexed with the help of bookmark option, and the overview of authors and illustrations is also available.



The second preservation method is archiving on external hard disks of great capacity ranging from 300 GB to 1 TB. External hard disks are used to record original digital copies in high resolution. These disks are used for protection only, not for re-writing.

The third preservation method uses servers which make digital copies available online. INFOBIRO archives data on two servers: one has all digital copies accessed by users, and the other is used to back up all data on server.

Conclusion

At present, INFOBIRO is the only organized digital archive in BiH. With the support of FBiH Ministry of Culture and Sports we are implementing the project "Digital preservation and archiving of cultural and historical heritage in BiH: BH press since 1866". Experience so far has shown that INFOBIRO has taken the right approach in presenting and making accessible a valuable content. The number of users visiting www.infobiro.ba site is about 30,000 a month and they come from about 72 different countries. Using PDF format in addition to

word text has proven to be a good choice. Converting from PDF into word format enables the users to search the text using their own key words. INFOBIRO does not index a content based on the subject, areas etc. as we think in this way users would be deprived of full information.

When processing newspapers of older date, INFOBIRO keeps the original language of the publication. When it comes to permanent preservation, research has shown that microfilming is still the most reliable option for permanent archiving. However, digital document is superior to microfilm when it comes to accessibility. In the future, the speedy development of IT will most likely produce media more reliable for permanent document archiving.

dragan@media.ba