

UDK (UDC): 004.9:004.738

Monika Péková*

DIGITIZATION OF ARCHIVES AND THEIR PRESENTATION ON THE INTERNET

Izvleček:

Digitalizacija arhivskega gradiva in njegova predstavitev na spletu

Masovna implementacija informacijske tehnologije na vseh področjih človekove aktivnosti vodi k preobrazbi v informacijsko družbo. Realnost teh, nekoč samo teoretičnih in filozofskih razmišljanj, je danes razvidna s pomočjo spletne tehnologije in novih informacijsko- ter komunikacijskotehnoloških storitev v splošnem. Digitalizacija arhivskega gradiva in njegova dostopnost v elektronski obliki predstavljata način, kako javnosti olajšati dostop. Članek obravnava prakso na področju digitalizacije arhivskih dokumentov ter rezultate in izkušnje projekta »Monasterium« pri digitalizaciji srednjeveških listin, ki jih hrani Slovaški narodni arhiv in so na spletu predstavljene preko portala Archival Vademecum.

Ključne besede:

arhivi, dostop do arhivskega gradiva, arhivsko gradivo, digitalizacija, izmenjava informacij, spletna stran.

Abstract:

Digitization of Archives and Their Presentation on the Internet

A massive introduction of information technology in every field of human activities leads to the transition to information society. Reality of these formerly only theoretical and philosophical considerations is presented nowadays by Internet technology and new Information and Communication Technology services in general. Digitization of archives and making them available in electronic form represents a way to facilitate the access for public. This paper discusses the practice in the field of digitization of archival documents, results and experiences of the project Monasterium, relating to digitization of medieval charters in the Slovak National Archives and presentation via the internet portal Archival Vademecum.

Key words:

archives, access to archives, archival document, digitization, information exchange, website.

A massive introduction of information technology in every field of human activities leads to the transition to information society. Reality of these formerly only theoretical considerations is nowadays presented by Internet technology and new Information and Communication Technology services in general. These possibilities create an infrastructure for a global information exchange and enable creation and use of information sources wherever it is and to anybody who has access to Internet. The most important influence of information systems on archives is created by the following facts:

* *Mag. Monika Péková, Ministry of Interior of the Slovak Republic, Department of Archives, Križkova 7, 81104 Bratislava, Slovak Republic.*

- possible digitization of different types of classical documents (text, image, sound),
- formation of hypertext, multimedial and hypermedial (www) documents,
- standardization of digitized documents and languages to organize character forms of www documents (SGML, HTML, etc.),
- standardization of exchange, communication and query protocols,
- possibilities of displaying and accessing different information sources containing information about documents and also full documents,
- possibilities of interconnecting different information sources and users wherever in the world.¹

The term digitization is very broad and includes several sub-activities and other related activities within the archival science. The designation "digitalization" in Archives can mean any of the following possibilities:

- conversion of finding aids into digital form;
- creation of digital copies of traditional records.

This paper focuses only on one topic in the area of digitalization in Archives: creation of digital copies of traditional archival documents and making them available to the general public.

Digitization of archival documents is the way, which allows access to the public through the use of digital copies and at the same time preserves original archival documents. Development of information technology in the last ten years has greatly influenced "life in archives". Archives have to cope with the transformation of the register of archival fonds and archival collections from paper form into electronic one, with creation of electronic records which later on become electronic archival documents.

The first consideration about digitization of archival documents in Slovak Republic was presented at the end of 2004. Between 2004 and 2006 the Department of Archives and Registries, Ministry of Interior of the Slovak Republic published some methodological directives.

It was supposed to build three digitizing centres in states archives - which have been established at the Slovak National Archives in Bratislava, the State Archives in Levoča and the State Central Mining Archives in Banská Štiavnica at the end of year 2006.

These centres have been equipped by scanners:

- digital equipment for scanning archival documents in format A2;
- scanner for microfilms;
- scanner Betterlight 6,000 for digitalizing large size and 3D documents;
- scanner Cruse for digitizing large format archival documents.

¹ Kimlička, Š.: *Knižnično-informačné systémy na prahu informačnej spoločnosti*. In: *Knižničná a informačná veda XVIII: zborník Filozofickej fakulty Univerzity Komenského*. Bratislava: Univerzita Komenského 1999, s. 7-31.

The main criteria for the selection of archival documents are their historic documentary value, scope of real physical damage and frequency of use by the research community.

Technical parameters for the aspect of digitization of archival documents stored in state archives provides the instruction issued by the Department of Archives and Registries Ministry of Interior of the Slovak Republic in November 2006. This instruction specifies the method of producing digital copies of archival documents (called safety copies) minimum in 300 dpi, TIFF format without compression.

Working versions are produced in 150 dpi, JPEG format, with 60 % compression for quick display. The instruction also settles the exact name of digital copy and information that must be contained in the XML file of the digital copy in order to be unambiguously identified.

Archiving - saving of the backup copies of scans are secured by their keeping in a high capacity storage system at the Slovak National Library in Martin and in respective Archives on DVD media. Inseparable part of digitizing is making information about archival documents and information from archival documents available to researchers by way of digital copies.

Information about archival fonds and archival collections administrated by state archives in the Slovak Republic is nowadays available via the internet portal Archival Vademecum, the first version of which was opened to the public in April 2008.² Archival Vademecum brings together in one central repository all relevant information about the archival fonds from the state archives.

The first part of the system is created by complex information about archival fonds and collections in the Slovak Republic and the data are taken from the nationwide archival information system. The archival information system is defined as a complex of information about archival documents recorded in the archival heritage register, about their storage, content and access to them. Nowadays the State Archives use a new version of electronic archival information system (AFondy).

Its aim is to create a central online database of the register of archival fonds and collections kept in the network of state archives as well as the register of archival fonds stored in all other types of archives. The database contains data about archival fonds and archival collections, their names, registration number, time extent, thematic group, amount extent, name of creator, information about physical status, state of processing and date of record updating.

The outputs from application AFondy are transferred to the data storage in order to be able to use them via a standard web browser. Through the common or full-text searching the user can quickly and easily obtain all available information from central data storage. If there exists a finding aid to a respective archival fond, Archival Vademecum enables to assign the information on finding aid or to present its full text.

The archival heritage of the Slovak Republic kept in state archives comprises more than 27,000 archival fonds and collections with a total range of more then 187,000 running metres.

Department of Archives at the Ministry of Interior of the Slovak Republic keeps the central register of finding aids. From the total amount of 4,000 units of finding

² Archival Vademecum, accessed on January 5, 2011 at <http://vademecum.vsnet.sk/sk/vade/>.

aids about 80% of them are in analogue / paper form and only 20% in digital form. The reason is obvious - the finding aids were made by typewriter until the 1990s and later on via personal computers. From year 2007 finding aids have been produced also in electronic form by means of the application ProArchiv - Inventories. These finding aids (more than 300 inventories) are accessible in electronic form and immediately provided on the Internet as we are aware of current trends and needs to provide information about archival fonds and archival collections through the Internet.

There are two possibilities to converse analogue finding aids into electronic form:

- manual transcription in the application ProArchiv - Inventory,
- digitizing from paper form into format PDF with OCR.

The conversion of all finding aids from paper into electronic form is the priority of Slovak archives as a preferable way of paper finding aids digitization. Inventories are digitalized into format PDF with OCR in co-operation with the Slovak National Library in Martin and should be soon provided on the respective website of Slovak archives.

The Slovak Act 395/2002 on Archives and Registries and on the Amendments of Certain Acts states: "The Archives shall only submit an original of an archival document for its study in exceptional cases, whilst it shall always submit a copy of an archival document, if it concerns an archival document created before 1526."³ This was one of the main reasons why the Slovak National Archives participated in the international project Monasterium.Net. The main objective of this project is to digitize all medieval charters in the archives, to make them available in electronic form for public use and preserve the most important and most historically valued archival documents - medieval charters from the Slovak National Archives that are important for the national and European history.

The medieval charters are administered by department of Early Fonds and Collections at the Slovak National Archives. Medieval charters by provenience are included in:

1. archival collections of ecclesiastical institutions,
2. a group of 98 archives of noble families and their feudal manors.

The digitization of more than 23,000 medieval charters began in May 2010 and ended in September 2010. The digitizing proceeded on two scanners which were provided by consortium Icarus. Medieval charters metadata are inserted in the application ProArchiv - Charters. This program contains twenty seven entries / items to be filled. For the time being only obligatory elementary items as Name of archives, Fond, Inventory number, Signature, Publisher, Receiver are filled in. Other data about medieval charters will be completing continuously later on.

Digital copies from all medieval charters stored in Slovak National Archives are available via a virtual guide through archival fonds and collections and their digital copies - Archival Vademecum.

The metadata from application ProArchiv - Charters will be transferred to the data storage in order to be able to use them via a standard web browser. Through

³ Article 12 of Act no. 395/2002 on Archives and Registries and on the Amendments of Certain Acts.

the common or full-text searching the user can quickly and easily obtain all available information from central data storage. The second part of the system Archival Vademecum represents the data gained from archival information system, which are refilled by information from finding aids produced in electronic form and the paper finding aids digitalized into format PDF with OCR.

The archives' repository makes the core of the entire solution. Structured image data in the JPEG format and metadata are stored in it. Above the repository there is an interface for the exchange of metadata using the OAI-PMH protocol. The OAI-PMH protocol enables harvesting the metadata automatically from the repositories of other data storage institution and in this way to share all relevant information sources. This function gives profit to users by offering the possibility to search for data across various local repositories. The viewing of digital copies will be enabled by the Zoomify tool which displays and zooms in on large images on the Internet without many demands on the user's hardware and Internet connection.

It is possible to summarize that the State Archives in the Slovak Republic are successful in gradually digitizing archival documents and making them accessible to the public. Important digitization projects were successfully launched in the recent past in spite of the limited budgets in the Archives. It is necessary to add that the results in terms of quantity and quality are mainly produced within the framework of international grant projects.