Digitization of Bulgaria's Ethnographic Archival Heritage

Ivaylo Parvanov, Elya Tsaneva
Institute for Ethnology and Folklore Studies with Ethnographic Museum
Bulgarian Academy of Sciences

Abstract: *The text was conceived as an account of the plans of the AIEFEM team to modernize, to make adequate to modern technical and scientific standards and to promote its archival heritage. Despite the poor conditions of preservation of the archival materials until now, the collections are enjoying significant interest and attention by different social and professional groups. The purpose of the described digitization is to enable the worldwide dissemination of valuable cultural visual data particularly of a preindustrial period and practices, less known as models of cultural manipulation.*

Keywords: scientific ethnographic archive, digitization, endangered collections
Ключови думи: научен етнографски архив, дигитализация, застрашени колекции

Ivailo Parvanov, PhD in Ethnochoreology, is an Assist. Prof. at the IEFSEM – Bulgarian Academy of Sciences, *working at the Department for Ethnomusicology and Ethnochoreology.*
E-mail: i_tparvanov@yahoo.com
Elya Tsaneva, PhD in Ethnography, PhD in Sociology, is an Assoc. Prof. at the IEFSEM – Bulgarian Academy of Sciences, *working at the Department for Scientific Information and Digitization.*
E-mail: etzaneva@gmail.com

INTRODUCTION

The Archive of the Institute of Ethnology and Folklore Studies at the Bulgarian Academy of Sciences in Sofia is a unique Bulgarian cultural collection – one of the oldest and definitely one of the most valuable. It consists of several separate collections of various materials: written documents (mainly fieldwork recordings, but also manuscripts, books, printed materials, maps, periodicals, watercolors, drawings, etc.); photo-, phono-, graphic, etc., documents, rich audio and visual collections. More than 15 years ago, it became clear to those working in the Scientific Ethnographic Archive that the current state of the collections was such as to necessitate their digitization; their present condition was considered liminal and critical with regard to preserving the informational and aesthetic value of the materials. Before becoming an important part of a huge digitization project for the establishment of a Center of Excellence – Heritage BG, the Ethnographic Archive was the subject of three earlier projects, implemented jointly with the British Library. This was in the years 2005 –2018 when Bulgaria had not yet deployed its own archival or library digitization program. That is why the Archive staff was proud to successfully use the capabilities of the

British Library and the Arcadia Fund in three programs for some truly life-saving work on unique archival materials. A brief overview of those project results, which were a precondition for inclusion in the current program, is presented as a first part of this report. Work on those projects is assessed to have been a source of early experience in digitization.

The first of these projects (EAP No103 2005 – 2006) resulted in the digitization of 3,865 glass plates; 1,911 old photographs and 4,684 negatives. Already in this first project, the team was trained in digitization and photographic techniques, gaining experience that was to be applied in the following years. After the successful completion of these activities, the Institute received a donation of materials from a private photo-collection, consisting of a total of 739 photo-units – black-and-white photographs and negatives that had been stored until then in extremely bad home conditions and were in need of urgent life-saving conservation work. The photographs illustrate elements of traditional culture: ritual masks, ritual breads, agricultural and animal husbandry activities, calendar customs, a complete collection of folk costumes and traditional architecture, etc. This new collection, together with the photographs selected by the team from the so-called "Old Albums" of the earliest Bulgarian photo studios (including the famous Karastoyanov studio, established in 1906 – 1907, and even earlier ones, such as photographs taken in 1886 by a private photographer), formed the content of Project EAP No 618 2012 – 2013. An interesting task of the project was the digitization of a collection called 'Folders with old negatives', dating from 1905 to the 1940s. The total number of negatives contained in them is 4,639. The main goal of project EAP No0989 2017 – 2018, for the first time, to digitize narratives and written documentary evidence – an extremely challenging task for the team members. In this connection, we decided to digitize three endangered manuscript collections from the 19th century, which we judged to be in imminent threat of physical destruction; they contained about 1,400 non-standard manuscript pages coming from a wide geographical area of Bulgarian and the Balkans, and related to elements of cultural and folk heritage. The first manuscript, entitled 'Antiquities. Throughout the Bulgarian Lands with Brief Descriptions', included two manuscripts, handwritten and hand-painted in ink, pencil and indigo by their collector and author Ivan Enchev – Vidyu; the work contains 470 pages plus 425 images inscribed in the text. The second collection consists of 525 handwritten pages which record cultural material dating from the early 20th century to 1927; this registration was the work of Konstantin Zagorov. The third collection includes 415 handwritten sheets of so-called short folklore forms, i.e., various genres of songs, proverbs, callings, riddles, etc., collected by Garnev. The total amount of digitized units was 1,416. This project of the Archive was in fact the first to include written surviews of the collections.

Today, the unique Ethnographic Archive is an important participant in a major digitization project, which includes the creation of a Centre of Excellence – the Heritage BG Consortium. Under this project, the team initiated a massive and sustained multi-source archival digitization program, in which the digitization process is conducted with advanced technology and ideology, as described below.

The main digitization activities of the project 'Preservation, presentation and socialization of the Bulgarian ethnographic and folklore cultural heritage (Preparation of a digital ethnographic map of Bulgaria)' are related to the digitization of different types of physical data carriers[1] from the archival and museum collections of IEFEM – BAS, and their integration (introduction, processing and visualization) into a specially developed computer-based information system (C – BIS)[2] 'Digital Ethnographic Map of Bulgaria'. Each

[1] In this article, the term "physical data medium" will mean any physical artifact (paper, polyvinyl chloride record, magnetic tape, videotape, film, museum exhibit, hard disk, etc.) that holds or contains data.

[2] "A computer-based information system is a designed system that encodes, stores, processes, and disseminates data about the state of an environment, using software and hardware solutions", see *Parvanov*, 2021: 27.

of these activities represents a separate stage of achieving the final goal set for the project team, namely the 'digital transformation' of the museum exposure through the methods of ethnographic mapping.

One of the main prerequisites for achieving this goal is the proper planning and implementation of the digitization stage. This stage is meant to ensure maximal quality of the digital copies of the physical media, leading to the sustainability of the exhibit material and a longer 'shelf life' of the final product. Hence, this article is entirely focused on the digitization process, considered in relation to the types of physical data media contained in the archival and museum collections of IEFEM – BAS.

1. THE TYPES OF DATA CARRIERS IN THE ARCHIVAL AND MUSEUM COLLECTIONS OF IEFEM – BAS IN RELATION TO THEIR DIGITIZATION.

Unique in content and diverse in form physical data carriers are stored in the archival and museum collections of the Department of Scientific Information and Documentation, the National Center for Intangible Cultural Heritage, and the National Ethnographic Museum at IEFEM – BAS. Depending on the technical means used for their digitization, they can be broadly divided into two main types:

- Physical data media, which can be digitized without the aid of an additional device for their reproduction (media whose data can be perceived directly by humans), such are data on paper and three-dimensional objects (museum exhibits);
- Physical data media whose digitization requires a reproducing device (media whose data can be perceived by humans in their entirety by means of a machine): data on audio-visual media (records, magnetic reels, cassettes, film and video tapes, records, etc.).

Most often used for digitizing the first type of media are scanners or digital cameras with various modes of operation[3], while digitization of the latter type requires a playback device – a "player" and an additional analog-to-digital converter.

2. THE DIGITIZATION PROCESS FROM A 'TECHNICAL' POINT OF VIEW

Digitization is 'the material process of converting discrete analog streams of information into digital bits'[4]. To clarify the technical side of the process, we will look at one of the most commonly used 'digitization' devices – the desktop scanner.

A 'scanner' is a peripheral device, 'that optically scans an image, printed text, handwritten text, or object, and converts it into a digital image'[5]. In terms of their principle of operation, scanners may be handheld, drum scanners, scanners with a movable scanning head, or three-dimensional (3D) scanners. The most common are those with a movable scanning head. In this case, the document to be scanned is placed so that the image faces a glass top (**Fig. 1**).

[3] For more information see: *Leggett* 2021; *Johnson* 2021; *Zigouris* 2021; *Butterworth, Pearson, Sutherland, Farquhar* 2018; *DIN – German national standard* 2019; *Iraci, Hess, Flak*, 2017; *Iraci*, 2017.
[4] *Brennen, Kreiss* 2014.
[5] *Pushkar, Sibilyev* 2011: 38.

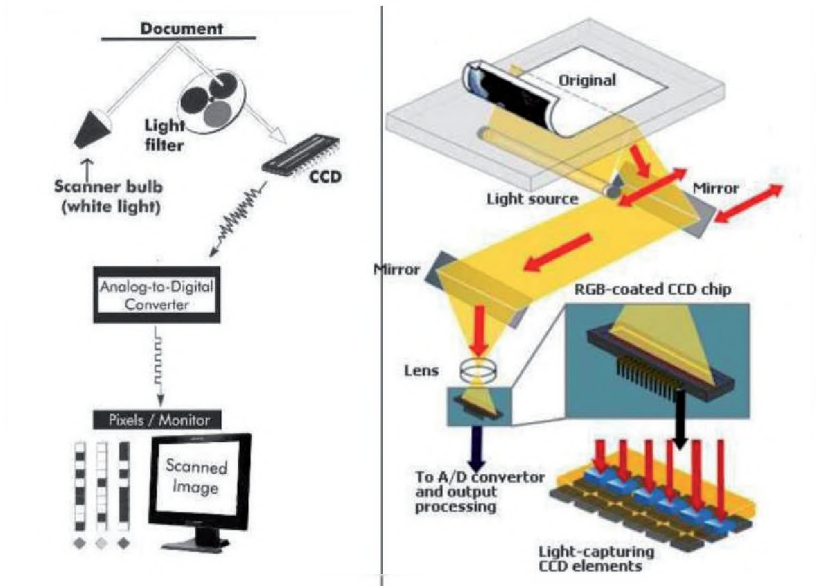

Figure 1. Scanner – working principle[6]
Фигура 1. Скенер – принцип на работа

After a user activates the 'scanning' process (pressing a button), a light fixture (usually a xenon or cold cathode fluorescent lamp) is triggered under the glass top, to illuminate a specific sector (row) of the scanned surface. By means of a mechanism composed of mirrors and lenses, the reflected light reaches a sensing and conversion device, the 'Charge-Coupled Device'[7], for the extraction of light images'[8]. Magnified, the surface of the 'CCD' looks like a large, dot-filled grid (matrix) (**Fig. 2**).

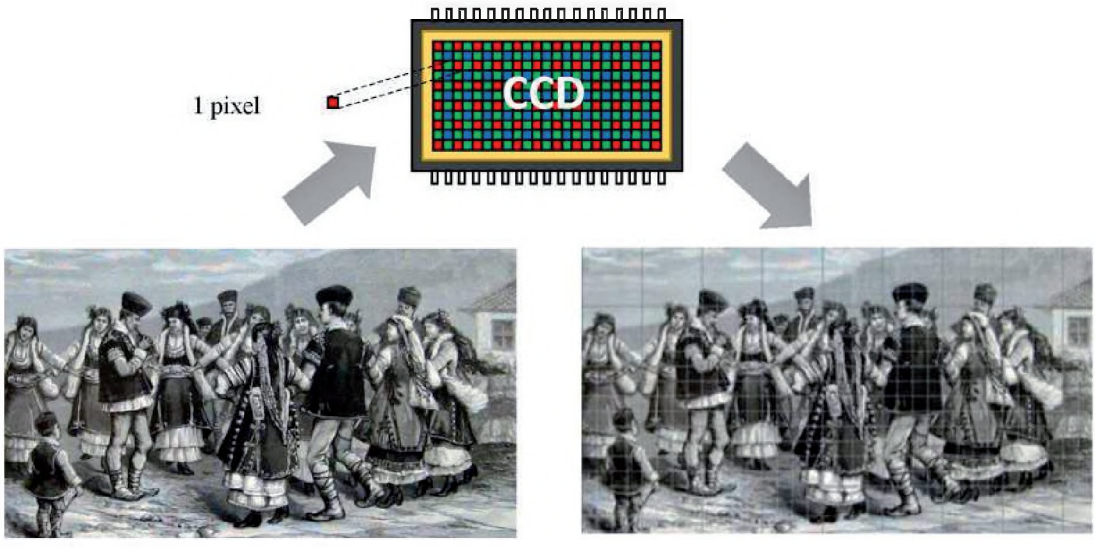

Figure 2. CCD – operating principle
Фигура 2. CCD – принцип на работа

[6] *Pushkar, Sibilyev* 2011: 39.
[7] Charge-Coupled Device (CCD) developed in 1969 by Willard Boyle and George E. Smith. The CCD is a special silicon, analog-integrated circuit that converts photons into electrical charges. I apply the abbreviation 'CCD' used in English publications.
[8] *Levski* 2017.

Each of these points is a light receptor called a photodiode. One dot (one photodiode) is equal to one pixel, which is the smallest constituent unit of a digital image. The larger the number of pixels (photodiodes), and the larger their size, the more detailed the image is. The total number of photodiodes (pixels) in a 'CCD' also determines its resolution, which is measured in 'dots per inch (dpi)'[9],[10]. Each photodiode (pixel), depending on the intensity of the light, converts the beam of light incident on it into an electrical signal[11] (a continuous-time signal), which an analog-to-digital converter, by means of an encoder, converts into a discrete-time electrical signal, referred to in the literature as a 'binary code'. The computer processes the sequence of electrical pulses ('binary numbers') that come from the ADC and stores them in their sequence of arrival in the computer memory. Thus, pixel by pixel, the picture on the paper is converted into a sequence of electrical pulses ('binary numbers'). This sequence, which is generally a rectangular or square matrix, is called a raster graphic. Each cell of the matrix contains a binary number that encodes the pixel color information[12]. At the end of the process, on the information display device, which in everyday life we call a monitor, we actually see a sequence of matrix cells that glow in different colors.

3. THE DIGITIZATION PROCESS FROM A 'HUMAN' PERSPECTIVE

As explained in the previous section, the activation of the 'scan' process (the pressing of a button) is performed by a human user. If we take a broader look at the digitization process, we clearly see that the actions performed by the human user (positioning the object on/in the scanning device, handling the mobile scanning device – handheld scanner, adjusting the settings in the software for computer processing of the signal received from the scanner, setting the camera, adjusting the lighting, adjusting the audio-visual playback device, etc.) are the basis for achieving maximal quality of the digital copy. Moreover, when the object of digitization are archival and museum collections, the importance of the archival and museum professional, who can assess the physical condition of the physical data media and their suitability for digitization, increases exponentially.

Therefore, the definition given by Brennan and Kreis, that digitization is 'the material process of converting individual analog streams of information into digital bits' should perhaps be reformulated, at least as far as digitization in the field of cultural heritage is concerned, into 'a human-managed material process of converting individual analog streams of information into digital bits'.

When putting the human factor in the focus of the digitalization process, we must inevitably take into account subjective factors like personal qualities, views and acquired knowledge and skills. Standards are set in taking these factors into account. As far as the digitization process is concerned, a standard for its implementation was published in 2010: *Guidelines for Implementing Digitization of Records*[13]. But this guide is for 'use in the design and conduct of responsible digitization by all organizations undertaking digitization, or business process digitization projects'[14] and is not fully applicable to the standardization of digitization related to cultural heritage. Far more specifically detailed on the subject are the various digitization guides of the leading institutions in the field, including National Archives of Australia, British Library, Smithsonian institute, Deutsche Digitale Bibliotek, Bibliothиque nationale de France, National Film Institute Hungary – Film Archive, Filmoteka Narodnowa, Fiaf, NFI, POD, and many others. The main problem of these guides is that they are designed to standardize work

[9] As in the case of TVs, the resolution of a 'CCD' is measured diagonally in inches.
[10] *Fulton* 1997-2019.
[11] *Stark* 1982.
[12] *Microsoft Docs* 2017.
[13] *ISO/TR* 13028:2010 Information and Documentation.
[14] *ISO* 2010.

in specific institutions with specific technical support (available equipment).

In consideration of all this, it was decided to draw on practical experience and, in view of the needs of IEFEM – BAS related to the process of digitization of archival and museum exhibits, to create our "internal" standard, which meets the requirements for our technical security and for the sustainability of the exhibit material (see **Tab. 1**).

4. THE DIGITIZATION PROCESS AS RELATED TO 'EXPOSURE'. THE ULTIMATE GOAL

Table 1. An example of a table for digitizing paper documents at A3 format
Таблица 1. Пример на таблица на дигитализиран хартиен документ в А3 формат

Format type	**Paper documents under A3**	
Purpose	Save	Access
File type	TIFF 6.0	JPEG
Compression	uncompressed	JPEG Compression Photoshop Level 10-12
Resolution	400 - 600 ppi	300 ppi
Bit depth	16-bit color (48-bit)	8-bit color (24-bit)
Colour space	sRGB or Adobe RGB	sRGB

As mentioned in the introduction to the article, the ultimate goal of the project team is the 'digital transformation' of the museum exhibit by means of ethnographic mapping. As for the digitization process, on the one hand it should ensure the best possible quality of the digital copies of physical data carriers, thus ensuring the sustainability of the exhibition material; on the other hand, it should provide a maximally informative digital presentation (exposure) of the physical data carriers. In other words, depending on the type of data medium, appropriate digitization methods must be applied to produce digital presentations that are as faithful as possible to the originals.

A considerable amount of practical experience has been acquired in work with paper-based data media and audio-visual data media; thus, their digitization and exposure does not present any difficulty. Museum exhibits, which, in the typology proposed in point 1, are three-dimensional physical data media, are a far more complicated matter.

Depending on the technical support and the type of physical data carrier, the methods used worldwide in the digitization of three-dimensional objects are 2D Photography, 360^{s} Photographic View, Laser Scanning, 3D Models by CAD (Computer Aided Design), Photogrammetry, CT Scanning, Structural Light Scanning and Touch Scanning. Each of these methods has its advantages and disadvantages, and each requires different kinds of financial resources and technical training of the digitization specialist. Taking into account all these factors, IEFEM – BAS, under the project 'NASLEDSTVO BG', purchased a photo laboratory (which predetermines the use of the methods 2D Photography and 360^{s} photographic view), as well as two 3D scanners for laser scanning of small-sized objects.

Falling within the scope of digitization carried out at IEFEM – BAS is a huge amount of traditional garments, which are essentially three-dimensional physical media. Their digitization and digital exposure are one of the most serious challenges for the project team. The not inconsiderable number of institutions (the Louvre Museum, the V&A Museum in London, the Kyoto Garment Institute, the Google Arts & Culture platform, etc.)[15] that deal with the digitization of textiles use the methods: 2D Photography, 360^s photographic view, 3D models produced by the extremely laborious CAD (Computer Aided Design) modeling and Photogrammetry. Considering the available technical resources at IEFEM – BAS, the team's efforts are currently focused on developing a proper approach to realizing 360s photographic views. The most basic problems that we encounter at this stage are: the different textures that make up a single traditional garment, the many layers of it that need to be visualized, the different sizes and genders of garments, which require different genders and sizes of mannequins, etc.

CONCLUSION

As the discussion above has shown, proper planning and standardization of the digitization process is a key point in the implementation of projects for digital preservation, presentation and socialization of the cultural and historical heritage. Depending on the end goal, the technical support, the type and condition of the physical data media, the expertise of the specialists, and the right management, sustainable results can be achieved that guarantee the maximum lifetime of the digital copies. The working model implemented at IEFEM – BAS, which was presented in the paper, reveals the challenges faced by any organization embarking on the not easy path of digital transformation. Challenges that very often boil down to choices and decision making against specific project frameworks. For us, they are not compromises, but a conscious compensatory mechanism that, with the help of a body of knowledge, aims to overcome the systematic neglect of the digitization process and the underfunding of heritage projects.

BIBLIOGRAPHY:

Brennen, Kreiss, 2014: *Brennen*, Scott; *Kreiss*, Daniel. Culture Digitally – Digitalization and Digitization.http://culturedigitally.org/2014/09/digitalization-and-digitization/ (05.03.2022).

Butterworth, Pearson, Sutherland, Farquhar 2018: *Butterworth*, Jody; *Pearson*, Andrew; *Sutherland*, Patrick; *Farquhar*, Adam. Remote Capture: Digitising Documentary Heritage in Challenging Locations. Cambridge, UK: Open Book Publishers.

DIN – German national standard 04.2019: DIN SPEC 15587 – Recommendations for digitization of cinematographic film. Retrieved from European Standards: https://www.en-standard.eu/din-spec-15587-recommendations-for-digitization-of-cinematographic-film/ (22.04.203).

Fulton 1997-2019: *Fulton*, Wayne. A few scanning tips. Retrieved on November 03, 2019, from https://www.scantips.com/chap3c.html (30.04.2023).

Iraci 05.2017: *Iraci*, Joe. The Digitization of VHS Video Tapes. Retrieved from Canadian Preservation Institute: https://publications.gc.ca/collections/collection_2018/pch/CH57-3-1-31-2016-eng.pdf (19.04.2023).

Iraci, Hess, Flak 05.2017: *Iraci*, Joe; *Hess*, Richard; *Flak*, Kimberley. The Digitization of Audio Tapes. Retrieved from Canadian Conservation Institute: The Digitization of Audio Tapes – Technical Bulletin 30 – Canada.ca (19.04.2023).

ISO/TR 13028:2010. Retrieved from International Organization for Standardization: https://www.iso.org/standard/52391.html (20.04.2023).

Johnson 02.2021: Preservation Digitization Standards for the Digitization of Physical RNA records. Retrieved from National Archives of Australia from https://www.naa.gov.au/sites/default/files/2022-01/Preservation-Digitisation-Standards-2021.pdf (20.04.2023).

Leggett 2021: *Leggett*, Elizabeth. Digitization and Digital Archiving: A Practical Guide for Librarian. Rowman & Littlefield.

Microsoft Docs. 2017: Types of Bitmaps. Retrieved on 10 11 2019 from https://docs.microsoft.com/en-us/dotnet/framework/winforms/advanced/types-of-bitmaps?view=netframework-4.7.2 (22.04.2023).

Pushkar, Sibilyev 2011: *Pushkar*, Oleksander; *Sibilyev*, Kostiantin. Information Systems and Technologies. Summary of Lectures Retrieved on October 12, 2019, from http://www.repository.hneu.edu.ua/handle/123456789/5879 (25.04.2023).

Stark 1982: *Stark*, Henry. Application of Optical Fourier Transforms. New York: Academic Press.

Villarreal.2020: The Use of 3D Apparel Simulation Software for Digitizing Historic. Retrieved from NC State University Libraries, PhD dissertation, manuscript, 2020: https://repository.lib.ncsu.edu/bitstream/handle/1840.20/38170/etd.pdf?sequence=1&isAllowed=y (22.04.2023).

Zigouris, 2021: *Zigouris,* Panagiotis. Contemporary Methods of Digitization. Athens: Panagiotis Zigouris.

Levski 2017: *Levski,* Deyan. Цифрови светлинни сензорни матрици – от ерата на динозаврите до днес [Tsifrovi svetlinni senzorni matritsi – ot erata na dinozavrite do dnes / Digital Photomatrices – from the Age of Dinosaurs to Today]. Balgarska nauka / Bulgarian Science (96).

Parvanov 2021: *Parvanov,* Ivaylo. Дигитален атлас на българския танцов фолклор. Зони на разпространение според теренните проучвания в етнохореологията. [Digitalen atlas na balgarskiya tantsov folklor. Zoni na razprostranenie spored terennite prouchvaniya v etnohoreologiyata./ A Digital Atlas of Bulgarian Dance Folklore. Zones of Dissemination according to Field Studies in Ethnochoreology.]. PhD dissertation, manuscript, Sofia: IEFSEM, 2021.

Дигитализация на българското етнографско архивно богатство

Ивайло Първанов, Еля Цанева

Текстът е замислен като описание на плановете и усилията на екипа на АЕИФЕМ да модернизира, да приведе в съответствие със съвременните технически и научни стандарти и да направи видимо наличното в него архивно наследство. Въпреки лошите условия на съхранение на архивните материали до момента, колекциите се радват на значителен интерес и внимание от страна на различни социални и професионални групи. Целта на представената дигитализация е да се даде възможност за световно разпространение на ценните визуални и наративни данни за културата, особено от прединдустриалния ѝ период, които досега не са били широко популяризирани.