

THE POWER OF INTELLIGENT CONTENT CURATION AND CONTEXT-BASED DIGITAL LIBRARY CONTENT USAGE FOR RESEARCH AND E-LEARNING (THE CASE OF NATIONAL LIBRARY “IVAN VAZOV” – PLOVDIV, BULGARIA)

Desislava Paneva-Marinova¹, Dimitar Minev², Ivan Krachanov², Detelin Luchev¹, Maxim Goynov¹, Lubomir Zlatkov¹, Radoslav Pavlov¹, Lilia Pavlova³

¹ *Institute of Mathematics and Informatics, Bulgarian Academy of Sciences (BULGARIA)*

² *National Library “Ivan Vazov” – Plovdiv (BULGARIA)*

³ *Laboratory of Telematics, Bulgarian Academy of Sciences (BULGARIA)*

CONTENT

This paper is focused on the digital library users' flexible and efficient access to multimedia representations of cultural and historical artifacts. It demonstrates the power of intelligent content curation and context-based DL content usage for research and e-learning. We present a prototype of a digital collection management environment, designed for the needs of National Library "Ivan Vazov" – Plovdiv. After a brief introduction, the digital archive of the library and the prototype are described, and finally possibilities for future developments are outlined.

- INTRODUCTION
- DIGITAL LIBRARY "IVAN VAZOV" – DIGITAL COLLECTIONS AND ARCHIVES
- DIGITAL LIBRARY "IVAN VAZOV" – STRUCTURES AND FUNCTIONALITY FOR INTELLIGENT CONTENT CURATION AND CONTEXT-BASED CONTENT USAGE FOR RESEARCH AND E-LEARNING
- CONCLUSIONS AND FURTHER DEVELOPMENTS

1. INTRODUCTION

- Digital libraries (DL) are powerful and efficient tools for content management (acquisition, storage, indexing, access, and maintenance), content enrichment, and structuring, as well as effective content search, filtering, and dissemination. Their rise over the past decade has given researchers and readers more opportunities to access, manage and organize the data they research or study. In this paper we present a technological solution that supports the National Library “Ivan Vazov” – Plovdiv in online storage, retrieval, curation, and observation of its va
- The prototype produced to manage the digital collections of the National Library “Ivan Vazov” – Plovdiv is a web-based software environment and includes not only basic modules such as a metadata management and presentation functional module (incl. specific services), a metadata model management module, administrative services, linked to the media repositories, and a user data repository, but also specific services, i.e. advanced search applied simultaneously in the metadata and the digital objects’ content; selection of library units through ordered or alphabetic lists or calendars; access to the environment through different devices, *etc.* luable cultural and historical artifacts.

1. INTRODUCTION

- The paper presents the descriptive structures for different objects used. Our focus is on the library users' flexible and efficient access to multimedia representations of cultural and historical artifacts, the support of rich functionality for content interaction and real-time adaptable and interactive end user centric visualization and its learning needs. In this study the power of intelligent content curation and context-based DL content usage for research and e-learning is demonstrated.

2. DIGITAL LIBRARY “IVAN VAZOV” – DIGITAL COLLECTIONS AND ARCHIVES

- The National Library “Ivan Vazov” in Plovdiv plays an essential role in the preservation of Bulgarian culture and history. It is the first cultural institution in Southern Bulgaria, established as a Regional Library and Museum of Eastern Rumelia in 1879, and has developed an archive of Bulgarian books and periodicals, a historical archive, a rich repository of manuscripts and Revival literature, and unique collections of rare and valuable publications.
- Today, the National Library “Ivan Vazov” in Plovdiv is a cultural institute that continues to dynamically enrich and develop the traditions of its prominent founders. Patrons annually realize over 120,000 visits and loan over 300,000 library documents. The library’s holdings are comprehensive and amount to over 1,900,000 library units – scientific, fiction, manuscripts, old-printed, rare and valuable publications, Bulgarian and foreign periodicals, photographs, maps, audiovisual and electronic documents, original works of art, personal libraries.

2. DIGITAL LIBRARY “IVAN VAZOV” – DIGITAL COLLECTIONS AND ARCHIVES

The digital archive of National Library “Ivan Vazov” offers nine collections:

- **BOOKS**, 194 titles currently available, predominantly Statues of professional organizations from Plovdiv and the region, from the end of the 19th century and the beginning of the 20th century. They are of interest to researchers because they provide insight into the way of life of the era.
- **PERIODICAL PUBLICATIONS**, a collection of newspapers and magazines currently containing 224 full-text titles, with approximately 23,000 separate issues. Particularly valuable are the newspapers and magazines from the East-Rumelian period: “Maritsa” newspaper - the first Bulgarian newspaper after the Liberation, “Narodniy glas”, "Nauka", "Zora", *etc.*, as well as the Revival-period collection of periodical publications.
- **MANUSCRIPTS**, a rich collection including Slavonic, Greek, Ottoman and Persian manuscripts on parchment and paper from the 11th to the 19th century.
- **GRAPHIC PUBLICATIONS**, a collection of art prints, lithographs, etchings, engravings, posters, original paintings, *etc.* The projects for monumental works are of special significance, with a total of 95 projects for large-scale wall murals, frescoes, ceramic tilework, most of them realized in many Bulgarian towns.

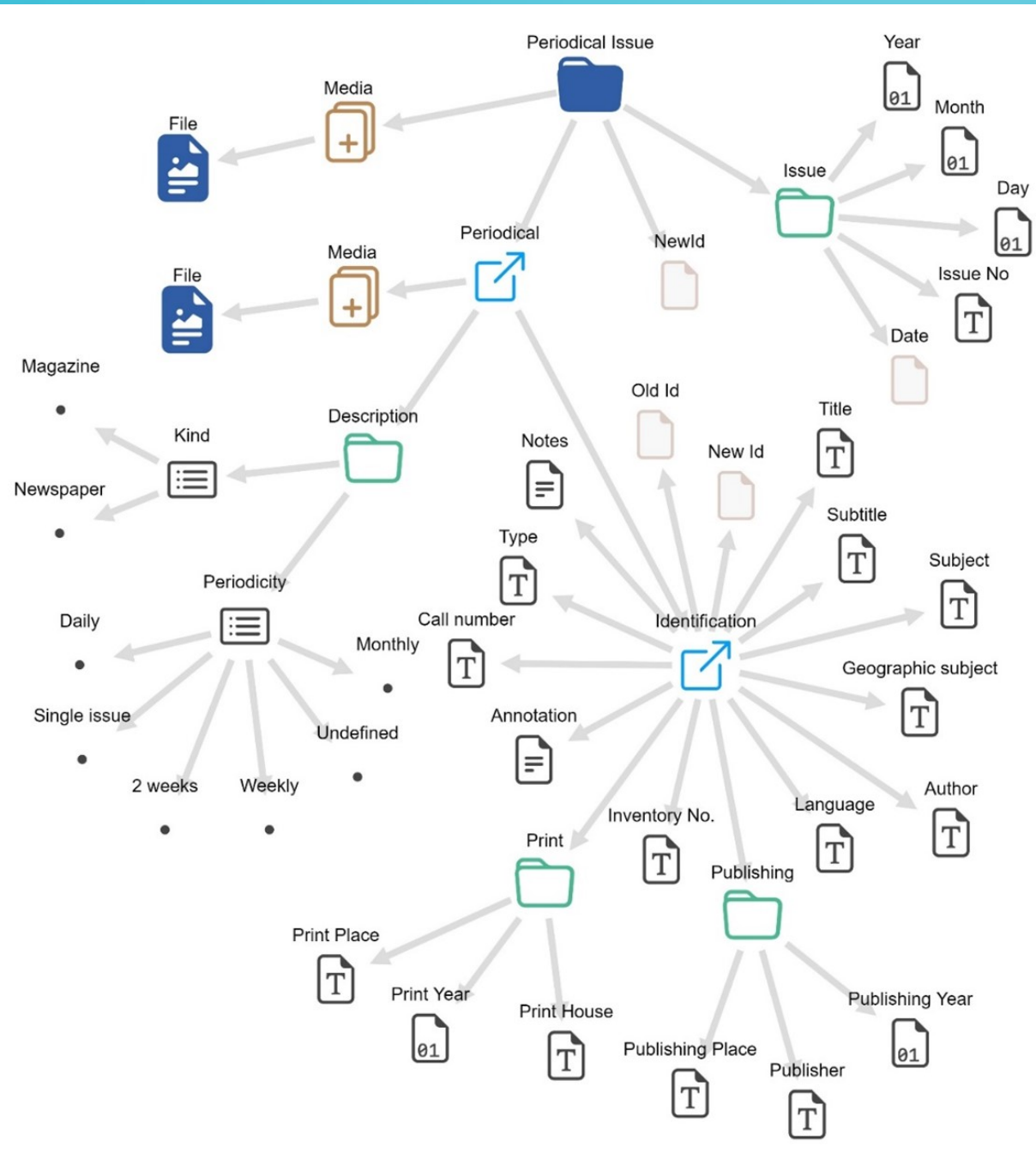
2. DIGITAL LIBRARY “IVAN VAZOV” – DIGITAL COLLECTIONS AND ARCHIVES

- **CARTOGRAPHIC PUBLICATIONS**, a collection of the most valuable possessions of the library, including the oldest map of Bulgaria, created by a Bulgarian and printed in Bulgarian language - "Map of the present Bulgaria, Thrace, Macedonia and the adjacent lands" from 1843.
- **PHOTOGRAPHS**, a digital collection currently consisting of 163 photographs. The library has a physical collection of approximately 4,000 photographs and postcards, portraits, and pictures of events and sites of historical significance.
- **ARCHIVES DIGITAL COLLECTION**, a collection of documents of nationwide value, revealing key moments of the political, economic and cultural development of Bulgaria. The documents cover the period from the 12th to the 20th centuries, the most numerous being from the second half of the 19th century.
- **AUDIOVISUAL DOCUMENTS**, the software platform of our Digital library allows us to publish audio and video file formats. The library has a rich collection of classic films on 35mm reels - masterpieces from the birth of cinema in 1895 to the early 40's of the 20th century. The library also has a large collection of recorded music with approximately 16,000 vinyl records.
- The library also has a rich collection of **SHEET MUSIC PUBLICATIONS**.

3. DIGITAL LIBRARY “IVAN VAZOV” – STRUCTURES AND FUNCTIONALITY FOR INTELLIGENT CONTENT CURATION AND CONTEXT-BASED CONTENT USAGE FOR RESEARCH AND E-LEARNING

3.1. Descriptive Structures for Efficient Access to the Literary Heritage Knowledge

- Specially created descriptive structures for different library objects are used for users' flexible and efficient access to their knowledge representations. The knowledge levels and the values of the metadata are determined after extensive analysis and scientific study of the essence of the presented objects and the respective fields of literary heritage from which they originate.
- Each library unit is digitalized and represented according to a descriptive model, based on the data and knowledge available. Figure 1 illustrates schematically the descriptive model for an object “Periodical issue” and its basic descriptors. The knowledge level “Issue” incorporates the temporal identification aspects for a single issue, such as issue number and year, month, day, and/or date of publication. The knowledge level “Periodical” relates to descriptors for the identification and description of the primary title.











	A	Array of objects (descriptors)
	B	A set of common descriptors
	C	Reference to model
	D	Single line text descriptor
	E	Multiline text descriptor
	F	Numeric value
	G	List of values
	H	Hidden element (internal use only)

Figure 1. Structure for object of type “Periodical Issue”

3. DIGITAL LIBRARY “IVAN VAZOV” – STRUCTURES AND FUNCTIONALITY FOR INTELLIGENT CONTENT CURATION AND CONTEXT-BASED CONTENT USAGE FOR RESEARCH AND E-LEARNING

3.1. Descriptive Structures for Efficient Access to the Literary Heritage Knowledge

- The structure implements the “one to many” relationship using arrays of elements (A). The “many to one” and “many to many” relationships are built using the model references (C). References are also used to avoid redundancy and keep data normalized. In some cases, denormalization is needed in order to improve search performance. It is implemented as a separate software layer, so it is not part of the conceptual modelling. D, E, F, G represent scalar data values. Hidden elements (H) are used to store data, which is not part of the software kernel, but is important for some internal cases like data migration, reference matching, custom search, indexing, *etc.*

3. DIGITAL LIBRARY “IVAN VAZOV” – STRUCTURES AND FUNCTIONALITY FOR INTELLIGENT CONTENT CURATION AND CONTEXT-BASED CONTENT USAGE FOR RESEARCH AND E-LEARNING

3.2. Intelligent Content Curation and Context-based Digital Library Content Usage for Research and e-Learning

- The basic functional components are in the core of the services for intelligent content curation and context-based digital library content usage for research and e-learning: a *metadata management and presentation functional module (incl. specific services)*, a *metadata model management module*, *administrative services* that are linked to a media repository and a user *data repository*. The basic digital library prototype stores and manages the digital analogues of cultural heritage objects presented in text (via .pdf files, fully corresponding to book media), graphic, video, audio formats, or other media objects as well as the relevant metadata. The module for management of the metadata model includes a service for building the descriptive schemas (descriptor structure) for cultural objects, the so-called model builder. The basic activities related to the creation and management of metadata for digital library objects: adding, storing, editing and deleting metadata; searching, selecting (filtering), accessing, viewing and displaying metadata, are implemented in the functional module for managing and presenting metadata.

3. DIGITAL LIBRARY “IVAN VAZOV” – STRUCTURES AND FUNCTIONALITY FOR INTELLIGENT CONTENT CURATION AND CONTEXT-BASED CONTENT USAGE FOR RESEARCH AND E-LEARNING

3.2. Intelligent Content Curation and Context-based Digital Library Content Usage for Research and e-Learning

- Specific services for intelligent content curation and context-based digital library content usage for research and e-learning, as a part of this module, include functionalities closely related to the content provider needs, such as advanced collection creation, management and curation (thematic collections, time dependent collections in a calendar structure, etc.), search in the text media objects, advanced objects preview and ordering, different device support, etc. The National Library “Ivan Vazov” – Plovdiv prototype provides specific services for indexing Portable Document Format (PDF) objects, providing the opportunities for full-text search in the objects’ content (Figure 2). Functionalities for presentation of PDF objects in a web environment (without browser add-ons requirements), with options for search and visualization of the results in a certain document, are integrated in the system.

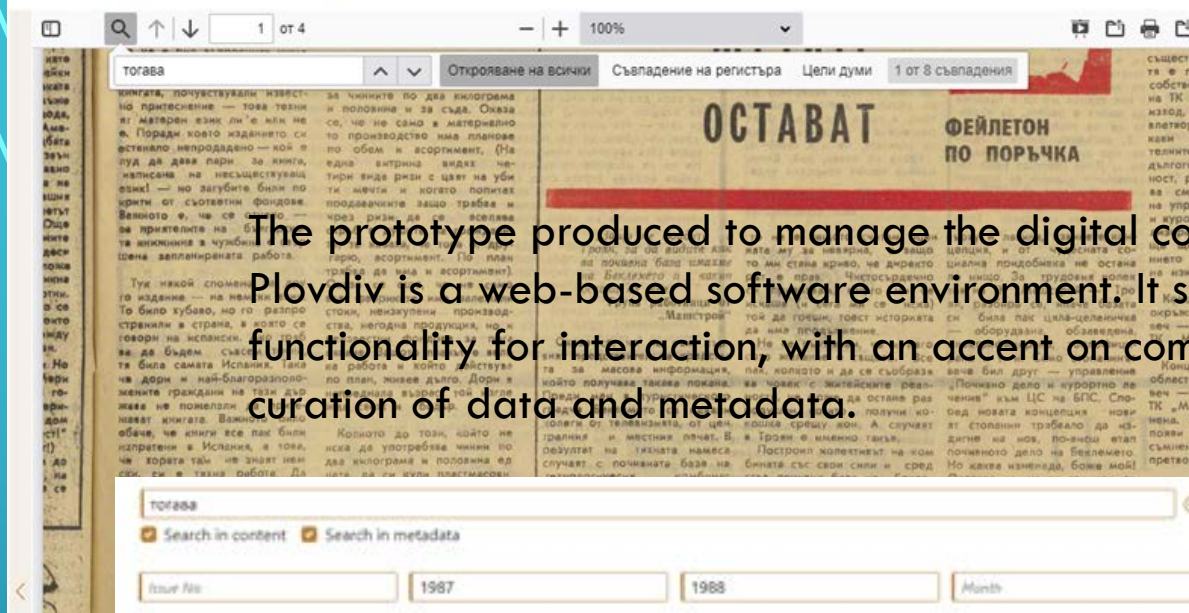
3. DIGITAL LIBRARY “IVAN VAZOV” – STRUCTURES AND FUNCTIONALITY FOR INTELLIGENT CONTENT CURATION AND CONTEXT-BASED CONTENT USAGE FOR RESEARCH AND E-LEARNING

3.2. Intelligent Content Curation and Context-based Digital Library Content Usage for Research and e-Learning

- The intelligent content curation and context-based digital library content usage for research and e-learning is supported by automatically implemented searching and tagging of a term (word form) throughout the database. The digital library provides advanced search applied simultaneously in the metadata and the digital objects' content (Figure 3).
- The full text search functionality is implemented not only for the text content, but for the objects' metadata. The library units could be selected through ordered or alphabetic lists (Figure 4) or calendars (Figure 5).

Figure 2. Full-text search in the digital objects' content

СТЪРШЕЛ , Брой № 2208, 03.06.1988



The prototype produced to manage the digital collections of the National Library “Ivan Vazov” – Plovdiv is a web-based software environment. It supports a variety of digital cultural units and rich functionality for interaction, with an accent on components providing storage, retrieval and intelligent curation of data and metadata.

Figure 3. Content search functionality applied to all editions from a year



Figure 4. Library units ordering

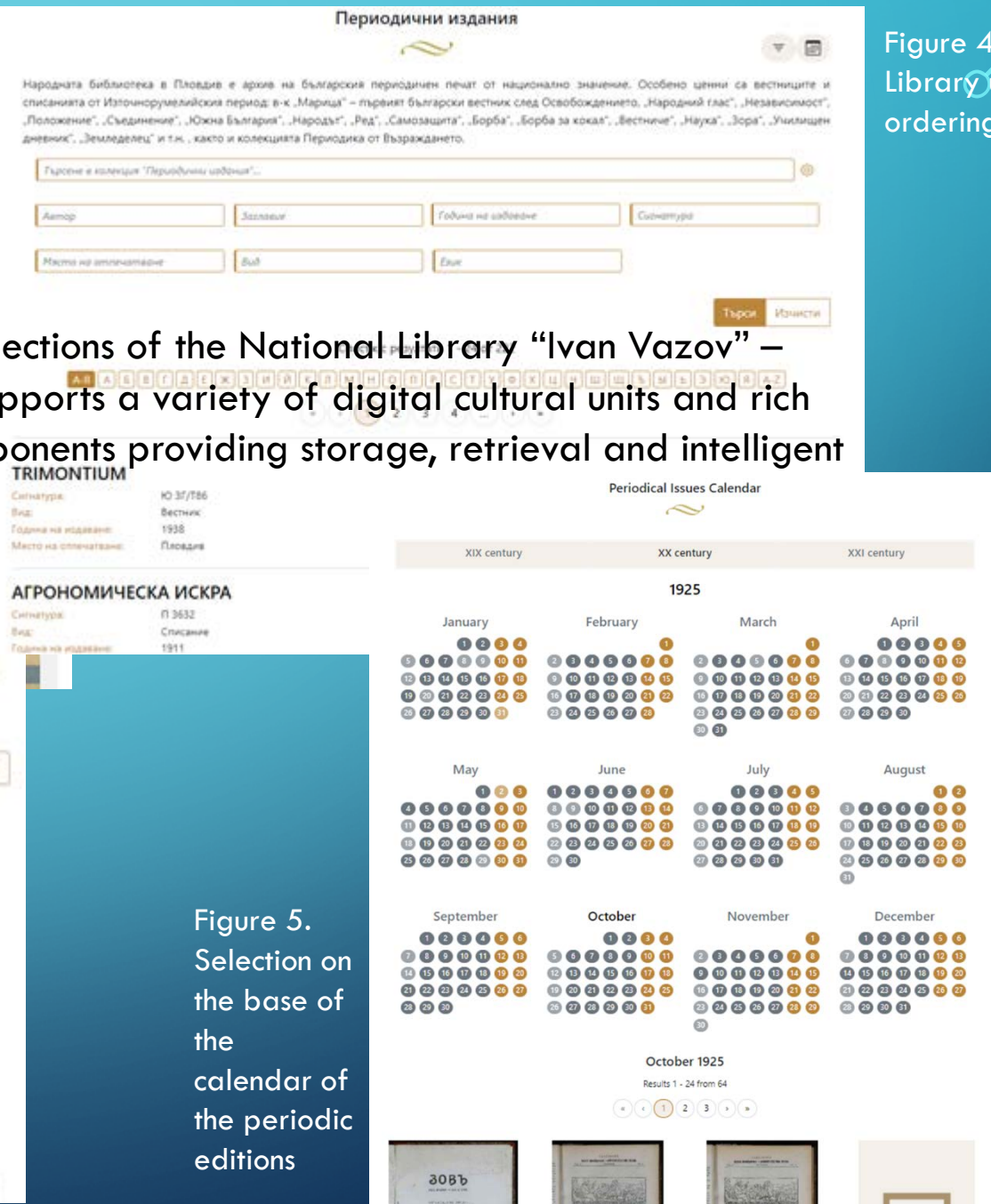


Figure 5. Selection on the base of the calendar of the periodic editions

3. DIGITAL LIBRARY “IVAN VAZOV” – STRUCTURES AND FUNCTIONALITY FOR INTELLIGENT CONTENT CURATION AND CONTEXT-BASED CONTENT USAGE FOR RESEARCH AND E-LEARNING

3.2. Intelligent Content Curation and Context-based Digital Library Content Usage for Research and e-Learning

- For wider accessibility and use of the digital library environment for research and e-learning, it can be accessed through different devices (Figure 6) – PC, smart TV, tablet, smart phone and a customized preview is supported. For the purposes of certain research and e-learning, a full text and a regular search in objects with different structure, but with some common characteristics, can be carried out (Figure 7).
- Supporting the context-based digital library content usage for research and e-learning, there are also options for placing links to pop-ups with text and media files (images), basic data analysis/synthesis, *etc.* Aggregation of objects is done based on their common (one or more) characteristics, depending on the specific model and application domain. The system summarizes groups of objects based on the aggregated data to improve the organization of object representation in subsequent analysis.

Figure 6. Preview on different devices

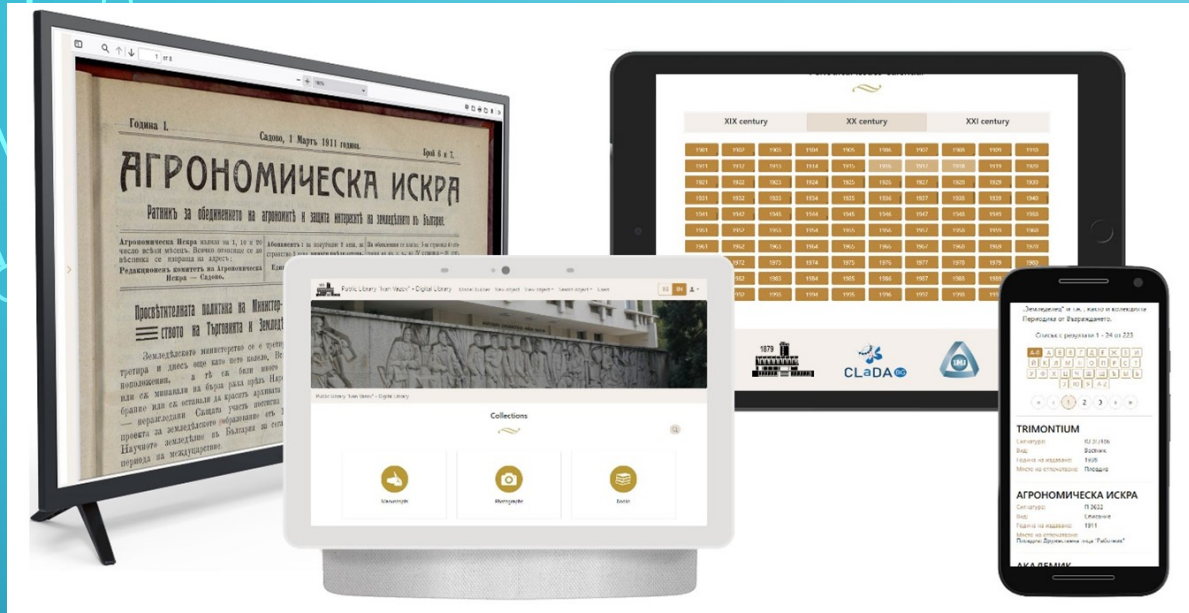
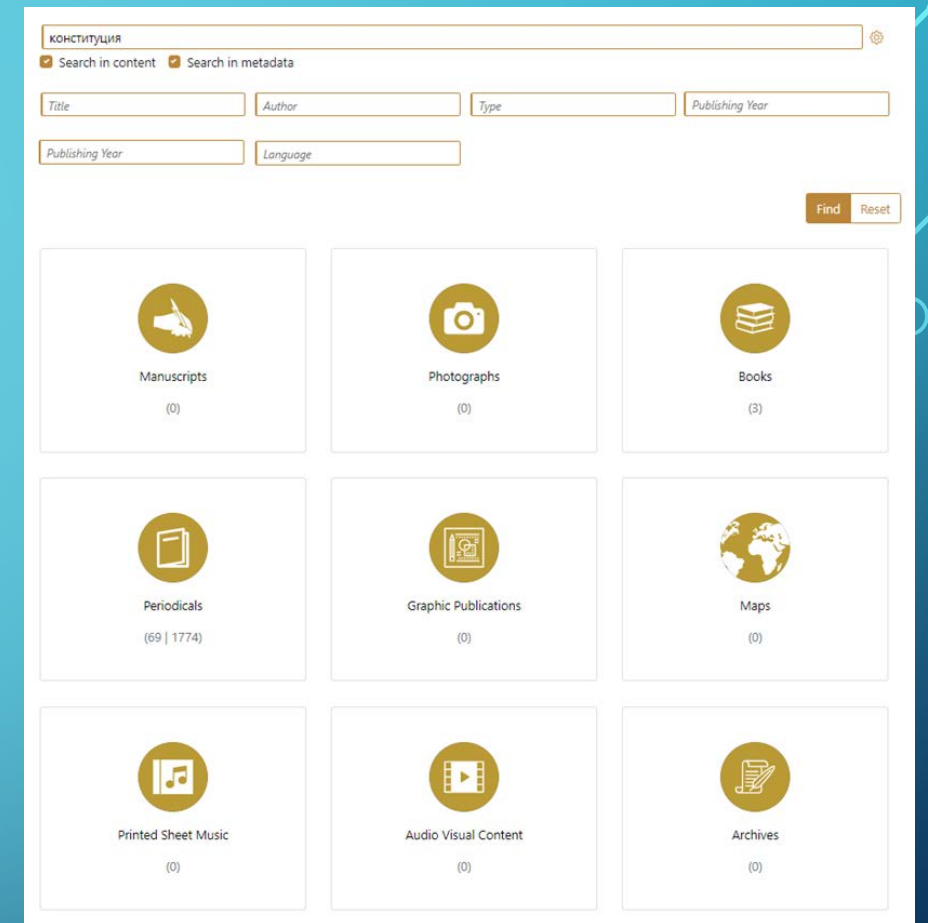


Figure 7. Full text and regular search in objects with different structure, but with some common characteristics



- Most of the digitised materials, those of high cultural and historical value and expired copyright in the collections of the National Library “Ivan Vazov” – Plovdiv, are predominantly from the period before the Bulgarian language’s last Orthographic Reform of 1945, which poses a number of problems to the accuracy of OCR and which the platform’s search algorithm must take into account in order to achieve results as accurate as possible. A very specific aspect of the platform’s search function is that it must take into account the peculiarities of optical character recognition (OCR) applied to Cyrillic texts. This provides additional opportunities for intelligent content curation and context-based digital library content usage for research and e-learning, as not all researchers and students are familiar with Bulgarian spelling before the Orthographic Reform of 1945.

4. CONCLUSIONS AND FURTHER DEVELOPMENTS

The prototype aims at providing the functionalities for maximizing the learning and research experience of the users of the digitalized archives of the National library “Ivan Vazov”. With this in mind, it could be further extended in a number of different directions, with the potential design and development of the following components (among others):

- additional functionalities for presentation of expositions, collection and exhibitions in virtual reality environment;
- a full-text index, based on dictionaries of written Bulgarian before the Orthographic Reform of 1945, providing more efficient search in digitalized books and periodicals (joint work with the National Library “Ivan Vazov”);
- a module for visualization of semantic nodes;
- integration of the data input forms with the Google Translate API (or similar service) and local functionality for transliteration from Cyrillic to Latin script;
- interactive tables in the admin panel, providing the tools for searching, filtering, sorting and design of specific data from particular object types in tabular form;
- formal description schemata and specific graphic interfaces for all collections of the National library “Ivan Vazov”;
- a functionality module for digital watermark protection of the content.

ACKNOWLEDGEMENTS

This work is funded in part by CLaDA-BG, the Bulgarian National Interdisciplinary Research e-Infrastructure for Resources and Technologies in Favor of the Bulgarian Language and Cultural Heritage, Part of the EU Infrastructures CLARIN and DARIAH, Grant number DO01-377/18.12.2020.

REFERENCES

D. Paneva-Marinova, R. Pavlov, K. Rangochev, “Digital Library for Bulgarian Traditional Culture and Folklore,” Proceedings of the 3rd International Conference dedicated on Digital Heritage (EuroMed 2010), 8-13 November 2010, Lymassol, Published by ARCHAEOLOGIA, pp. 167-172, 2010.

D. Paneva-Marinova, M. Goynov, D. Luchev, Multimedia Digital Library: Constructive Block in Ecosystems for Digital Cultural Assets. Basic Functionality and Services. LAP LAMBERT Academic Publishing, Berlin, Germany, 2017.

D. Paneva-Marinova, J. Stoikov, L. Pavlova, D. Luchev, “System Architecture and Intelligent Data Curation of Virtual Museum for Ancient History,” SPIIRAS Proceedings, vol. 18, no. 2, pp. 444-470, 2019.

R. Stewart, M. Zheleva-Monova, Y. Zhelev, L. Pavlova, D. Luchev, D. Paneva-Marinova, R. Pavlov, “The Orthodox Icons Collection of the Regional Historical Museum—Burgas: A Digital Library for Iconographic Objects,” Digital Presentation and Preservation of Cultural and Scientific Heritage, vol. 5, pp. 157-172, 2015.

D. Luchev, M. Goynov, D. Paneva-Marinova, J. Stoykov, L. Pavlova, “Synergy of National Cultural Heritage and Technology,” Digital Presentation and Preservation of Cultural and Scientific Heritage, vol. 11, pp. 281-286, 2021.

The background is a blue gradient with white circuit-like lines in the corners. The lines consist of straight segments connected by right-angle turns, ending in small circles, resembling a PCB layout or a network diagram.

Thank you for your attention!