

Stefana Janićijević

(Matematički institut SANU,
Beograd, Serbia)

**INSTITUTIONS IN DEVELOPMENT PROCESS
OF DIGITIZATION HERITAGE WITH STANDARDS, POLICIES
AND STRATEGIES IN SERBIA
CASE STUDY: MATHEMATICAL DIGITAL LIBRARY**

Abstract. We present the main of the achieved results in the field of digitization in Serbia. As introduction, the author analyzes the management of digital preservation and digital contents and figures approaches and actions briefly. Following this part, global situation in Europe in digitization and particularly the situation in Serbia within projects and institutions is presented. The final part is about s specific sort of digital library – mathematical digital library, about its sense, purpose and reason of existence. Several institutions and several projects developing are presented.

The value of scientific heritage in the knowledge treasury

Preserving of Serbian collected knowledge from scientific and cultural fields has been transformed during the periods according to technological achievements. Papers like this are often used to describe analyses about heritage integration, but the fact is that, from long ago in the past, culture heritage as a field has been theoretized in a majority of projects, as well as in the eyes of those determining priorities at the general level. It is governmental policy (Ministry of science and Ministry of culture, primarily) that set up global heritage systematization and standardization because public officials from state institutions are decision makers in *treasury of knowledge management*.

Heritage documentation today is at a transitional stage. On one side the advancement of technology, which is reflected by the diversification of topics and issues, has improved methodologies and data integration processes. On the other side there is a more widespread recognition of documentation as being an integral part of the conservation process, and not just an extra item to do if time and money allow it.

These responsible Ministries develop a complex sphere of interlinked rules and interests, policies, recommendations and directives. When entering this sphere, it is difficult to imagine that any sector or any segment of culture and science can be ignored. Even if they have not been directly regulating some fields, Serbian heritage society have been going through profound changes and everyone who lives and works in that environment is affected. This is particularly evident in the multidiscipline institutions which have been going through many changes and reforms, according to adoption of the European Union standardization. It is important to say that

the heritage area standardization in Serbia still needs to be remodeled and systematized by different actors and different levels in the decision making policy.

It is true that certain initiatives by public, private or civic sectors have been undertaken in order to create a good platform for establishing system in the digitization policy. Researching of the authors and projects in the field of digitization we notice the need for revitalization inherited institutional heritage system, lack of systematic approach in standardization and not existing of stable homogeneous policy in cultural heritage institutions. However, the classification of cultural and scientific heritage into the entities of World Heritage departments would refresh information and metadata about the Serbian legacy.

Many of local institutions have prejudice that digitalization is the best solution for all existing problems with heritage storage. Libraries, museums, archives, galleries and other similar institutions, public and private, follow contemporary trends from Europe and USA. On one hand, digitization represents basic instrument in achieving the goal of introducing inter-local, international and diversified culture, knowledge and science, which are aspects of intercultural dialogue. Still a large majority of institutions which deal with preservation of cultural and scientific heritage in Serbia are far from using digital modeling and a data base network with on line catalogue of institutions has not yet been developed.

Here are the facts common to almost all projects related with digitization process in Serbia:

- First projects in the field of digitization were started in the early nineties of the last century and they started as an initiative at University level
- Many of the projects were established with the principal aim of presentation
- The principal institutions participating in these projects were the Mathematical Institute of the Serbian Academy of Sciences and Arts, National Library of Serbia, National Museum of Serbia, the Faculty of Mathematics, the Archeological Institute of the Serbian Academy of Sciences and Art and the City Museum of Belgrade. Other institutions participated occasionally, e.g. the Musicology Institute of the Serbian Academy of Sciences and Art
- The greatest part of funds for financing these projects were coming from government sources, mainly from the Ministry of Science and Technology of Serbia
- Digitization in Serbia is part of new attractive idea of transforming our cultural policy and identity, and Ministry of Culture in the Republic of Serbia set up digitization as the main goal in heritage production
- Many state institutions (cultural and scientific) adopt digitization in their policy papers
- Several private and NGO institutions started with digitization activity, but it has not yet been implemented in market as a profitable process

Digitization in Europe and Serbia

Standards:

- EAD (Encoded Archival Description) is a nonproprietary encoding standard for machine-readable finding aids such as inventories, registers, indexes, and other documents created by archives, libraries, museums, and manuscript repositories to support the use of their holdings

- DTD (Document Type Definition), is the formal set of rules that define the elements that may occur within an encoded document and their structural relationships (their relative order and nesting)
- EAC (Encoded Archival Context) is an XML standard for encoding information about the creators of archival materials and the circumstances of record creation and use. It is intended to integrate EAC into Encoded Archival Description for enhancement of EAD's capabilities in encoding finding aids. EAC is defined in a document type definition (DTD) as well as in a XML Schema and a Relax NG scheme
- METS (Metadata Encoding and Transmission Standard) is an XML-compliant standard for encoding a variety of metadata about digital library objects
- DCMI (Dublin Core Metadata Initiative) is an organization dedicated to promoting the widespread adoption of interoperable metadata standards and developing specialized metadata vocabularies for describing resources that enable more intelligent information discovery systems
- Learning Object Metadata is description of resources for knowledge management. Main standard is IEEE for Learning Object Metadata (1484.12.1-2002.)
- MARC (Machine-Readable Cataloging) is the initial standard for catalog information retrieving and description of electronic elements resources. Anglo American Cataloging Rules adopted it.
- MODS (Metadata Object Description Schema) is a XML scheme for explanation of the technical, administrative and descriptive metadata
- VRA (Visual Resources Association) Core Categories Version 3 is scheme for description of art, architecture, artifacts and other visual resources. Visual Resources Association adopted it
- MPEG-7, Multimedia Content Description Interface is multimedia description and indexing of system which combine XML content with textual indexing of physical parameters
- SMPTE Metadata Dictionary is standard for audio visual formats developed by Society of Motion Picture i Television Engineers
- SGML (Standard Generalized Markup Language, ISO 8879) emerged from the analysis as being a technique able to meet all of the functional requirements as well as one supported by a large and growing number of software products available for a variety of operating systems. SGML is a set of rules for defining and expressing the logical structure of documents thereby enabling software products to control the searching, retrieval, and structured display of those documents
- TEI (Text Encoding Initiative) is an international consortium that publishes the TEI Guidelines for Electronic Text Encoding and Interchange, an SGML- and XML-compliant encoding language for the capture of literary and linguistic texts, widely used in the scholarly and cultural heritage communities
- Z39.50 is a standard which specifies a client/server based information retrieval protocol. It was originally proposed (in 1984) for use with bibliographic information. On the server, several data sources may be available. A client is able to search the server and the server returns results to the client
- DjVu is a document standard for efficient storing the scanned paper documents that could include color graphics, photographs in relatively small files suitable for web publishing
- SRU is protocol for Internet information retrieving and it uses CQL8 (language for queries). Library of Congress adopted it

- OAI (The Open Archives Initiative) is an attempt to build a "low-barrier interoperability framework" for archives (institutional repositories) containing digital content (digital libraries). It allows people (Service Providers) to harvest metadata (from Data Providers). These metadata is used to provide "value-added services", often by combining different data sets
- National Information Standards Organization (NISO)
- American National Standards Institute (ANSI)
- International Organization for Standardization (ISO)
- ISO/TC46/SC4
- ISO/TC46/SC9

Technology aspects:

Standards for text processing and graphic processing:

The digitization standards are by no means connected to metadata. Among all standards probably the most important are resolution of scanned documents, file formats for images and file format in which the digitized manuscript is kept. There are still debates among archivists what should be the resolution of scanned sources. It varies between 300 dpi and 600 dpi.

Most of scans are 600 dpi, some of them are 400 dpi, and the early ones are 300 dpi. Digitized works on display in most libraries are compressed files in 300 or 400 dpi; the uncompressed scans of higher resolution are kept for the archive purpose and possible future processing.

Particular scans are kept in tiff or pcx bitmapped file format. As already stated, digitized manuscripts are presented in pdf or DjVu file format.

Fedora system is a programming platform for administrating and accessing the database of digitized manuscripts. Fedora is a system specifically designed for handling digitized values, and among other benefits, it is open source software. Pandora software system is an archival description system.

Technologies are creating cheap and fast process of preparing digital collection for presentation, possibility of automatic generation of presentation from various formats of metadata, but with no additional conversion from online version to offline version.

Languages, formats and tools:

- HTML, DHTML, XML, XSLT, DOM, DDT, XHTML, XSL-F, XML-Shema, SVG (language for describing two-dimensional graphics in XML), JavaScript, Semantic Web, Web 2.0, ASP i PHP technology, ORM (Object Role Modeling), CSS, ECMA Script, SMIL (Synchronized Multimedia Integration Language), Unified Modeling Language (UML), TeX, OCR (Optical Character Recognition)
- Picture and media formats - JPEG, GIFF, TIFF, AVI, WAV, MPEG3, MPEG4
- Windows i Linux operation systems, Internet browsers – Internet Explorer, Netscape, Opera, Mozilla, etc.

Europe

Documents. In the light of regional mapping, it is important to be said about digitization *pioneer* project in EU: A group of national experts and representatives held an exploratory meeting in Luxembourg in November, 2000. This was followed by a meeting in Lund, Sweden, in April 2001, of Ministry of Culture representatives of member states and the European Commission. The Lund meeting recognized the importance of European cultural content and agreed on a series of principles which established a commitment to the creation and management of European cultural e-content.

Principles:

- Established the value of online European cultural content
- Identified barriers to realization of the full potential of the content
- Agreed steps that should be taken by Member States to overcome these barriers
- Outlined the role of the European Commission in these endeavors, by stimulating good practice and improved organization, as well as by supporting research

While these principles were agreed and endorsed by the meeting, it was perceived that a concrete set of Actions were needed in order to begin the implementation of these Principles. A follow-up document was drafted and agreed, which set out a series of steps to be undertaken. These are the *Lund Action Plan*. In November 2004 the Council of the European Union agreed to proceed with the coordination of digitization through an updated European action plan as a follow up of the Lund Action Plan, to be delivered under UK Presidency. The Netherlands, Luxemburg and UK Presidencies drafted the revised plan and its final version is the *Dynamic Action Plan for the EU coordination of digitization of cultural scientific content* as per 27th October 2005. The Dynamic Action Plan addresses the challenges of coordinating digitization activities to build a European cultural information space that is accessible and usable by citizens in a period of rapid technical development.

Six objectives are pursued through this updated action plan, acknowledging and building upon the previous set of Lund Principles:

- Providing strategic leadership in a dynamic and changing environment in which rapid technological and economical developments are taking place
- Strengthening coordination and forging stronger links between Member States' digitization initiatives, EU networks and projects
- Continuing efforts in overcoming fragmentation and duplication of digitization activities and maximizing synergy
- Assessing and identifying appropriate models, funding and policy approaches to sustain development and long-term preservation strategies
- Promoting cultural and linguistic diversity through digital content creation
- Improving online access to European cultural content

Various coordination actions, funded under the European Commission and European Commission's Information Society Technologies (IST) Framework Program (FP). Their aim is to provide policy makers and professionals working in scientific and cultural institutions at the local level (in particular, institutes, universities, public libraries, museums and archives) with a concise and relevant summary of the state of art in the use of new technologies.

Different digital guidelines manuals (terms) about standards, issues, practice in digitization policy of decision making.

Serbia

Projects and institutions:

- The Southeast European Digitization Initiative (SEEDI) is initiative which develops awareness about digitization of cultural and scientific heritage in the South-Eastern European countries along the Lund Principles of the European Union. It will contribute to gather and spread specific and interdisciplinary knowledge from various institutions in the region and European Union with leading experts in the field work.
- National Center for Digitization (NCD) is initiative and consortium includes Mathematical Institute and Archaeological Institute of Serbian Academy of Science and Art, Mathematical Faculty in Belgrade, National Library of Serbia, National Museum in Belgrade, Archive of the Republic of Serbia and Serbian Institute for Monument Protection and Yugoslav Film Archive. NCD officials proposed strategy for the next years: *Recommendation for coordination of digitization of cultural heritage in South Eastern Europe*.

The main subjects of NCD cooperation are the following:

1. Coordination of efforts of institutions involved in the cultural and scientific heritage digitization
2. Establishing and promoting a national strategy for the cultural and scientific heritage digitization
3. Exploring, adaptation and implementation of international standards and protocols for the cultural and scientific heritage digitization and preservation at the national level, development of new standards in areas where they do not exist
4. Launching the cultural and scientific heritage digitization and making plans for possible migration process to new formats and technologies for already digitized data

In 2002, NCD established electronic journal, called *Review of the National Center for Digitization*. Several conferences under the name *New Technologies and Standards: Digitization of National Heritage* were organized since 2002.

- One more digitization project in Serbia is project National Digital Library (NDL), in scope of National Library of Serbia (NLS). National Library of Serbia has been developing a Digital Library since 2003. After these four years, Digital National Library of Serbia presents a unique resource of digitized cultural heritage in Serbia and closer environment, providing free and full access to various digital collections and to half million of digital documents. It was built on the principles of an open access to the knowledge and information and offers the Serbian national treasure online. The final achievement of the NLS is a subscription of the contract for association in World Digital Library. The contract is drawn in Washington D.C., in Library of Congress, year 2008.

- Archiving the journal Publications de l'Institut Mathématique – Mathematical Institute SASA¹
- Collected works of Bogdan Gavrilović – Mathematical Institute SASA
- The Digitization of the Serbian PhD Theses – Mathematical Institute SASA
- Digital card of cultural heritage of Toplica Area, GIS technology – Mathematical Institute SASA
- Geographical Card Digitization with Geographic Information System (GIS) – Mathematical Institute SASA
- Kragujevac Journal of Mathematics – Prirodno-matematički fakultet Kragujevac
- Matematički vesnik – Društvo matematičara Srbije
- The teaching of mathematics – Društvo matematičara Srbije
- Visual Mathematics, e-journal – Mathematical Institute SASA
- Memorial compact disk – Faculty of Mathematics in Belgrade
- Bulletin, Classe des Sciences Mathématiques et Naturelles, Sciences mathématiques – SASA
- Applicable Analysis and Discrete Mathematics – Publications of the Faculty of Electrical Engineering Series Mathematics and Physics
- JAIS – Archive of the Republic of Serbia
- Development of data base – Cinemateque of Yugoslavia
- Viminacium – Archeological Institute SASA
- Vinča – Center for Digital Archeology, Faculty of Philosophy
- 3D Reconstruction of Gamzigrad – Center for Digital Archeology, Faculty of Philosophy
- Interactive Belgrade – Center for Digital Archeology, Faculty of Philosophy
- Mammoth – Center for Digital Archeology, Faculty of Philosophy
- Collection of engravings – City Museum of Belgrade
- Collection of Groman's old photographs – City Museum of Belgrade
- Old maps, engravings and photographs – City Museum of Belgrade
- Collection of photos of Jeremija Stanojević – City Museum of Belgrade
- Computer archiving and multimedia presentation of cultural values and national heritage, The project lasted for two years (1995–96) – Mathematical Institute, Mathematical Faculty, Archeological institute, and City Museum of Belgrade
- Nikola Tesla Museum Clipping Library, the archive of Nikola Tesla Museum contains the clipping library with thousands of clippings from newspapers and other publications. The library contains original clippings and many Tesla's annotations –Nikola Tesla Museum
- Digitalization of the Archive Resources of Nikola Tesla Museum, old photographs, technical drawings, etc. – Nikola Tesla Museum
- Digitalization of the Local Collection Public Library of Kruševac started with non book materials from local history collections – Public Library of Kruševac

¹ SASA – Serbian Academy of Sciences and Arts

- Star Maps Digitization and Connection with Data Bases, project presents old astronomical star maps as the part of world cultural and scientific heritage – Mathematical Institute SASA
- Project of Digitalization and processing of Archive Resources in Dubrovnik – SASA Archive
- Information system on stationary cultural goods – Serbian Institute for Monument Protection
- Catalogue of cultural monuments – Serbian Institute for Monument Protection
- Multiplatform data base for documentation of conservation process – Department for preservation *Diana* of National Museum in Belgrade
- One of very few NGO's which performs the continuous process of digitization is the Internet Club in Ljig. It pertains to a relatively small community of a population of 4,000 near Belgrade. This organization of enthusiasts works mostly with libraries, and its web portal www.biblioteke.org.yu is the only network of public libraries in Serbia

Case study –Digital Mathematics Library (DML)

One of the main tasks of the nowadays mathematical libraries is to provide integrated access to mathematics both on classical library resources as well as on heterogeneous and distributed digital information systems. What is needed is some kind of decentralized but integrated access.

International Mathematical Union, American Mathematical Society and European Mathematical Society have created a project about developing Digital Mathematics Library a concerted international collaborative effort to digitize the bulk of the world's mathematical literature, in the context of a world-wide endeavor endorsed by national mathematical societies in Europe and by the International Mathematical Union. Therefore, one of the first step is digitalization of global relevant mathematical heritage, as part of cultural and scientific heritage. It is about 50 millions pages with mathematical text, which implies that this project is of global world significance.

The European Mathematical Society seeks the support of the European Union in contributing to the Digital Mathematics Library (DML). The goal is to provide the framework for the creation of the world digital mathematical library and to put the European mathematical community at the leading edge of this creation, to make the digital mathematical library as comprehensive as possible concerning the mathematical literature produced in Europe and make the usage of this library available to the mathematicians in the world. The use of the old mathematical results, documented in the past mathematical literature has always been extremely important for the creation of new results and for the application in other fields, in natural sciences and engineering as well as in developing of technologies.

This digital mathematical library will be much more affordable for companies and university libraries. Easy access will be essential. The digital mathematics library will be monitored by the stake holders through the governing body of the World Mathematical Library. In this body the European Mathematical Society and its corporate members will be represented. Mathematics on digital systems includes digitized mathematics, digitally represented mathematics, and formalized mathematics. Moreover, there are different collections and projects working, partly independent, on the vision of a global Digital Mathematics Library. For many scientists, in particular in non-experimental sciences like mathematics, the Internet is becoming the main me-

dium for the exchange and storage of scientific knowledge. At the same time, the Internet becomes essential in the organization of scientific institutions such as research centers or international scientific societies. Therefore, there is an effort for Digital Mathematics Library which is a project coordinated by Cornell University Library and funded by the U.S. National Science Foundation (NSF) toward the establishment of a comprehensive, international, distributed collection of digital information and published knowledge in mathematics.

The Virtual Mathematics Library – VML (other name of DML) is projected as a cooperative, distributed, hybrid library. That means it will offer a structured access to important electronic resources as well as to the extensive library resources of the special subject collections. It is intended to become a central access to information and related services which are relevant for mathematicians. This could be realized in a web-based portal system. The concept of the VML includes remote access both to the contents as well as to the services of the corresponding information sources. It could provide access to and delivery from global mathematical knowledge systems.

The origin and the location of the project for a Virtual Mathematical Library is the State and University Library Göttingen, SUB Göttingen. Some of the institutions which also develop this project: Mathematical Institute of SASA (Belgrade), Österreichische Zentralbibliothek für Physik, Réseau National des Bibliothèques de Mathématiques (France), Swiss consortia of libraries, Deutsche Mathematische Vereinigung, London Mathematical Society, Société Mathématique de France, Société de Mathématiques Appliquées et Industrielles. Also, there are many projects related in digitization of mathematics materials and those are mostly connected with different national initiatives (like project that is correlated with National Center for Digitization in Serbia).

It could be created procedure for implementation mathematical digital library into European Digital Library (EDL) by:

Preparing roadmaps and incentives for mathematical institutions to bring existing and newly digitized material into EDL, encouraging private content holders to make their copyrighted material searchable and accessible through the common multilingual access point, agreements or collective agreements between right holders and scientific institutions, such as institutes, archives, libraries or museums, by which the latter can make copyrighted material accessible online on contractual terms.

To reinforce national strategies and targets for digitization and digital preservation by:

- Drawing up and updating plans and national strategies for digitization of mathematical material
- Establishing national strategies for long term preservation and deposit
- Developing quantitative and qualitative targets including the associated financial planning on a multi annual
- Basis for deposit, digitization and online access of scientific material and long-term preservation
- Investigating, and where appropriate initiating and promoting public-private partnerships for digitization

Selected projects:

- Project Eram provides a digital archive of the most important mathematical publications of the period 1868–1942 and a database based on the *Jahrbuch über die Fortschritte der Mathematik*

- The Euler–portal belongs to mathematics publications and the NSF/DFG Cornell–Michigan project for a distributed library of mathematical monographs as outstanding examples
- Virtual Mathematical Library is intended as a cooperative, distributed, hybrid library. The project is applied for funding through the DFG in the frame of the Digital Library Initiative, University Library Göttingen, SUB Göttingen
- Math Lib Page project is specialized access to university libraries, which is standardized for mathematicians
- Limes–Large Infrastructure in Mathematics Enhanced Services - project achieves the upgrade of the existing database Zentralblatt-MATH into a European based world class database for mathematics (pure and applied), by a process of technical improvement and wide Europeanisation
- Fedora The Open–Source Digital Repository Management System - Fedora Commons is a non-profit organization providing sustainable technologies to create, manage, publish, share and preserve digital content as a basis for intellectual, organizational, scientific and cultural heritage by bringing two communities together
- The Digital Mathematics Archive is a digital collection of mathematical sources, with a primary focus on documents from the late 19th century through today. Papers, letters, manuscripts, e-mail messages, Usenet-news postings, computer rendered images and computations, program source code, many of today's source documents are much more ephemeral than the documents of the previous century. The University of British Columbia Mathematics Department develops it
- Emani - the aim of the Electronic Mathematics Archives Network Initiative (Emani) is to provide a long-term preservation and access to mathematics literature. A network of reference libraries will provide the infrastructure. Borne digital content will be supplemented with retrospectively digitised material
- Numdam - the Numdamweb site (Numérisation de documents anciens mathématiques) offers an open access to the metadata and articles published in mathematical journals
- Active math - the ActiveMath group works at the frontiers of e-Learning. Group enables learners to tap their full potential by keeping up their motivation and fostering a self-regulated learning process. Group cooperates with teachers, tutors, psychologists and educators to improve learning and teaching in schools, universities and life-long learning
- Monet - the aim of the MONET project is to demonstrate the applicability of the latest ideas for creating a semantic web to the world of mathematical software, using sophisticated algorithms to match the characteristics of a problem to the advertised capabilities of available services and then invoking the chosen services through a standard mechanism

Bibliography

- Giuliana De Francesco, *Towards and agreed European Common Platform for digitization of cultural and scientific heritage: The Ministerial Network for Valorising Activities in Digitisation (MINERVA)*, Review of NCD 6 (2005), 2–11.
- Katharina Habermann, *A Project for a Virtual Mathematical Library*, New Developments in Electronic Publishing AMS/SMM Special Session, Houston, May 2004 ECM4 Satellite Conference, Stockholm, June 2004, pp. 75–80.

- Mile Jovanov, *System for storing data about national heritage with advanced search techniques*, Review of NCD 9 (2006), 53–58.
- Nebojša Vesić and Tamara Butigan-Vučaj, *All platform digital collections presentation*, Review of the NCD, 2006.
- *Recommendations for coordination of digitization of cultural heritage in south eastern Europe*, Review of NCD 7 (2005), 2–7.
- *Some Best Practices for Retrodigitization*, Committee on Electronic Information and Communication June 29, 2005.
- Žarko Mijajlović, Zoran Ognjanović, *A survey of certain digitization projects in Serbia*, Review of NCD 4 (2004), 52–61.
- Žarko Mijajlović, Zoran Ognjanović, Aleksandar Pejović, *Digitization of mathematical editions in Serbia*, Review of NCD (2007).
- Žarko Mijajlović, Zoran Ognjanović, Nada Đorđević i Tijana Zečević, *Virtual library - data base of textual data*, Review of NCD (2004).
- Žarko Mijajlović, Zoran Ognjanović, Aleksandar Pejović, *Digitization of Scientific Journals in Serbia*, Proceedings ELPUB2007 Conference on Electronic Publishing, Vienna, Austria, June 2007.
- Electronic editions of Serbian mathematical journals, http://www.mi.sanu.ac.yu/biblioteka/elect_pub.htm
- Electronic Library of Mathematics, <http://www.emis.mi.sanu.ac.yu/EMIS/ELibM.html>
- Encoded Archival Description, <http://www.loc.gov/ead/>
- Faculty of Mathematics, Belgrade, <http://www.matf.bg.ac.yu>
- EDL Project, <http://www.edlproject.eu/>
- Minerva Projects, <http://www.minervaeurope.org/>
- NCD, Belgrade, <http://www.ncd.matf.bg.ac.yu>
- Virtual library of old mathematical works, <http://alas.matf.bg.ac.yu/biblioteka/home.jsp>
- The European Library, <http://www.theeuropeanlibrary.org/>
- The Dublin Core Metadata Initiative, <http://dublincore.org/>