Zoran Krstulović, Mojca Šavnik National and University Library Ljubljana, Slovenia

LIBRARY MATERIAL DIGITISATION PROJECT COORDINATION AND E-CONTENT NATIONAL AGGREGATOR IN THE FIELD OF CULTURE

Abstract: In 2007, the Government of the Republic of Slovenia designated in the Strategy of Information Society Development in the Republic of Slovenia – si2010 the vision and defined the goal of e-content development - to ensure access to the e-contents in the Slovenian language in all fields that significantly influence the life and work of an individual. In the field of culture accelerated digitisation of existing analogue content, inclusion into the European digital cultural space and further development of the Digital Library of Slovenia have been especially stressed. In 2006, the National and University Library (NUK) established the Digital Library Development and Strategy Division, which manages all digital library processes and procedures. The article describes the basic tasks and objectives of the Division, including regular digitisation of library material and digital collections' metadata description. Recently, a new objective has been developed as a result of the needs of a wider Slovenian library community – i.e. co-ordination of digitisation in the Slovenian libraries. Lately, the established possibility of a deeper cooperation between the key actors in the fields of culture and science and the desire to coordinate the e-content creation have been reflected in the decision by the Ministry of Culture to establish the national econtent aggregator in the field of culture. The Ministry has appointed the National and University Library in Ljubljana, Slovenia as the national e-content aggregator. The European digital library Europeana defines aggregator as »an organization that gathers metadata from many cultural content providers and channels them into Europeana«. In the project outline, the purpose of the national aggregator development in Slovenia is defined more widely than in the Europeana context, i.e. »to enable a single point of web access to the Slovenian econtent in the field of culture, and to begin with the establishment of conditions for long-term preservation of econtent in the fields of culture and science«.

Key words: digitisation, coordination of digitisation, National and University Library, good practice, European aggregator, national aggregator, cultural heritage, repository of digital content, Digital Library of Slovenia

1. Introduction

The Digital Library Development and Strategy Division (Division) leads, develops and manages all processes of digital library in National and University Library (NUK). In addition to the existing functions of the Division, in the year 2010 under the Division has formed a new task, as the result or reflection of the needs in the Slovenian library environment. Paper includes coordinating the digitisation on the state level – namely: the preparation of technical documentation to carry out digitisation, collaboration with selected outsourced contractors for digitisation, monitoring of digitisation process, the final quality acquisition of digitised material and publishing on the portal of the Digital Library of Slovenia (www.dlib.si). Taking into account international standards, recommendations and guidelines for digitisation and building digital libraries, more than 30 projects of digitisation materials from different Slovenian libraries had been successfully carried out within the Division in 2010 – in addition to realiza-

tion annual digitisation plan of NUK for the year 2010. In addition to the realization of planed digitisation for NUK, the coordination of planned digitisation of library materials at the national level had started. That was an attempt to prevent duplication of digitisation of library materials and contribution to the rational use of public funds and virtual unification of library material that is spread over Slovenian libraries and beyond. Expressed need for national coordination of digitisation of library materials and the decision of the Ministry of Culture to establish a national aggregator of cultural e-content overlap and result in a successful project established a functioning national aggregator, which was awarded the National and University Library in the implementation and management.

2. Digitisation of NUK library materials

Division carries out an annual program of digitisation of NUK library materials following the parameters of the guidelines for the digitisation of library materials [1]. Entire NUK is involved in preparing the annual digitisation work program. All suggestions and wishes for

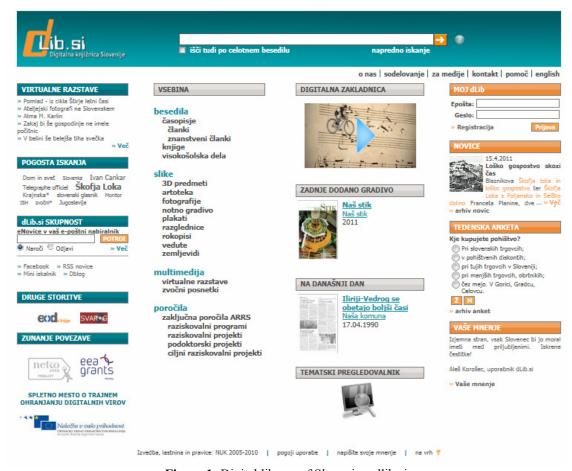


Figure 1: Digital library of Slovenia – dlib.si

digitisation are gathered in the Division, where the final version of the annual plan is also prepared. [2] General guidelines are a starting point and guide for the preparation of detailed technical documentation (TD) for the digitisation of various types of library materials. All digitisation of NUK Annual Plan is carried out through public tenders. NUK prepared in February 2008 an international tender for the first phase of the procedure with the preliminary determination of qualifications for the selection of candidates for carrying out the digitisation

of text book material [3]. There were four outsourced contractors selected for digitisation for four years. All selected contractors signed the Framework Agreement [4]. In early 2010 participation agreement with one of the selected contractor was discontinued. From 2008 until today digitisation is performed with known contractors, which over the years and projects had undertaken additional training in complex and specific digitisation of library materials.

NUK is the owner of the Digital Library of Slovenia portal [5]. Also, NUK manages, develops and maintains the portal. All NUK digitised library materials, which are obtained by carrying out digitisation projects in Division, are published on the portal dlib.si. Since the library materials are diverse and extremely challenging for the digitisation and also to publish the result online, it was necessary to adopt precise and explicit instructions for creating bibliographic and technical metadata, formats and file size.

2.1 Good practice. NUK with knowledge and ways of working, and international involvement in digitisation projects together with selected outsourced contractors of digitisation, developed the foundations of good practice. Good practice in the implementation of digitisation projects developed and confirmed in 2010 – good practice was confirmed more than effective and appropriate, as only four employees in one year managed to prepare with external contractors 1.2 million high-quality scans of NUK documents and published them in suitable format on the Digital Library of Slovenia portal.

The realization of the annual program of digitisation is usually initiated at the end of February, early March for the current year. The materials, prior to publication of tenders, are divided into thematic sections by type of material (monographs, serials, visual material ...) or medium (paper, microfilm). While preparing the tender an implementation plan of the project is accurately prepared also. The plan consists of technical documentation and timelines. Preparation of TD takes a long time, since the material, which is the subject of digitisation, has to be listed in, detailed and inventoried. The contractor is required to produce test digitised copies by the specification in TD and sent them to client for review and approval. Based on testing parameters certified copies the contractor carries out all digitisation. At the end of the digitisation NUK conducts quality review and a final quantity acquisition. With the approved final quantitative acquisition the project is completed for the contractor, while NUK has to publish digitised material on the portal dlib.si.

Technical documents consist of three thematic components: an inventory and description of the material, instructions for the creation of thumbnails and scans files, bibliographic and technical metadata, and instructions for proper handling of library materials during transportation and digitisation. An inventory and description of the material is inventoried: the number of pages, volumes, issues and condition of the material (brittle paper, radiation of letters through paper, difficulty opening the volumes ...). Resolution, colour depth, encoding and file name are determined in the second set. The contractor must prepare bibliographical and technical metadata according to the instructions, which are also part of the second set.

TEHNIČNA DOKUMENTACIJA

Digitalizacija listov Slovencev izven meja RS: Slovenski Primorec in Ilirski

Dokument vsebuje:

- Podatki o publikacijah
 Parametri digitalizacije in kreiranje metapodatkov
 Ravnanje z gradivom

1. PODATKI O PUBLIKACIJI:

naslov	št. letnikov	št. strani	višina strani	cobiss_id
Slovenski Primorec: podučen i kratkočasen list	1	4	41 cm	125599488
Seznam gradiva				
leto izida	letnik	številka	opombe in s	topnja ohranjenosti
1868	1	2	občutljivo;	slabem stanju in izredno evilka (nevezano)

naslov	št. letnikov	št. strani	višina strani	cobiss_id
Ilirski Primorjan: podučiven in omikaven list za slovensko ljudstvo	1	100	31 cm	125598976
Seznam gradiva				
leto izida	letnik	številka	opombe in s	topnja ohranjenosti
1866	_		gradivo je v	slabem stanju in izredno

1-24

občutljivo

1

2. PARAMETRI DIGITALIZACIJE IN KREIRANJE METAPODATKOV

Storitev

- 2.1 Digitalizacija serijskih publikacij 2.2 Optična prepoznava besedila (OCR) 2.3 Strukturiranje posameznih številk, povezovanje datotek, skenogramov in poimenovanje datotek

2.1 Digitalizacija publikacije

Parametri digitalizacije			
Primarni arhiv	marni arhiv		
Sivinsko			
ločljivost skeniranja	300 dpi		
barvna globina	8 bitov		
format	JPG		
Ogledna kopija			
ločljivost	96 - 150 dpi		
barvna globina	8 bitov		
format	JPG, PDF, HTML, TXT		

Ogledna kopija:
Skenogame primarnega arhiva naj se pretvori v ogledno kopijo z resolucijo 96 - 150 dpi in shrani v datoteki pdf formata. Pri tem naj bo nad visokoresolucijskimi datotekami izvedena optična prepoznava besedila (OCR). Rezultat je pdf datoteka, po kateri je moč iskati.

Formati datotek			
jpg datoteke	Skenogrami naj bodo v resoluciji, določeni s parametri digitalizacije. Ime .jng scan datoteke je lahko enako imenu .pdf datoteke. Enako ime mora biti tudi v xml shemi. Datoteke morajo biti: - obrezane (crop); - velike enotno število slikovnih pik; - poravnane.		
pdf datoteke	Datoteke naj bodo v ogledni resoluciji med 96 in 150 dpi (prilagojeno izvirniku). Pdf datoteke (multipage) naj se tvorijo za bibliografsko enoto - celotno številko publikacije.		
html	Besedilo mora biti brez oblikovanja.		



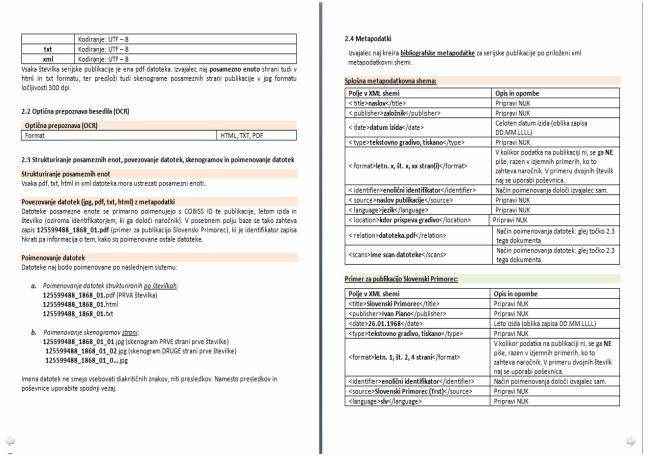


Figure 2: Example of technical documentation

3. Coordination of digitisation of library materials of Slovenian libraries

In 2010, spontaneously appeared the first need for national coordination of digitisation. NUK has already noticed in 2009 duplication of digitisation of library materials (Figure 3) in Slovenian libraries. Therefore the expressed need for coordination was most welcome among library community. Coordination was first conducted only in the form of data collection and information on the intended digitisation of the 2010 annual plans of libraries.

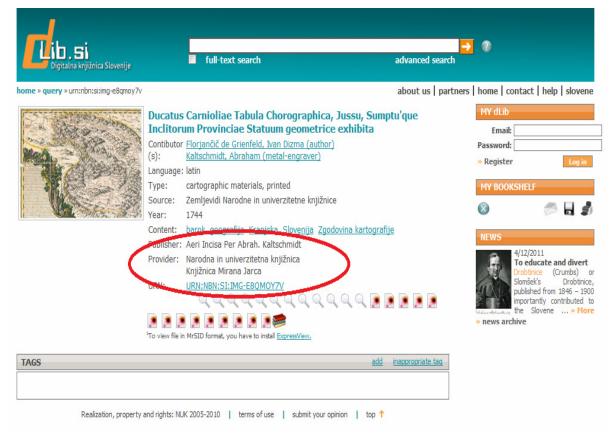


Figure 3: Example of duplication of digitising

Already in 2009 began publishing of digitised library materials from Slovenian libraries on the Digital Library of Slovenia portal. Starting in 2010, the Ministry of Culture determined that the digitisation of all library materials, which is financed from public funds, has to be published on the portal dlib.si. The complexity of the portal requires a detailed consideration of the technical parameters of digitisation and preparation of digitised material for publication on the portal. Libraries have begun to refer to the Division for help in preparing and implementing digitisation projects. Employees in the Division were initially only advising, and then were actively drawn into the projects. Libraries prepared lists of material, in Division we prepared a TD on the principle of good practice, which was by then already introduced to our range of activities. Such a system of work was possible because of sustained acquisition of Slovenica in NUK – because of that we have access to any material that was ever published in the Slovenian language and/or in Slovenia. We reviewed, surveyed and described the material, and at the client's request, also borrowed for the digitisation project. Rental of material was more the exception than the rule. Most of the material was rent to supplement the missing material.

Libraries are choosing among the selected NUK outsourced contractors for digitisation. This proved to be a very good solution. Contractors had already been trained to digitise complex library materials, so to do all work together to avoid beginners' difficulties and delays. Digitisation projects are fully managed in the Division – we confirm quality of test copies, advise and assist with problems and make the final quality acquisition of digitised materials. Projects are completed, when the material is published on the portal dlib.si. Only in 2010 there were 30 projects finished in the unit in such a way. Each library can find their own digitised material on the portal dlib.si by searching provider's name. (Figure 4).

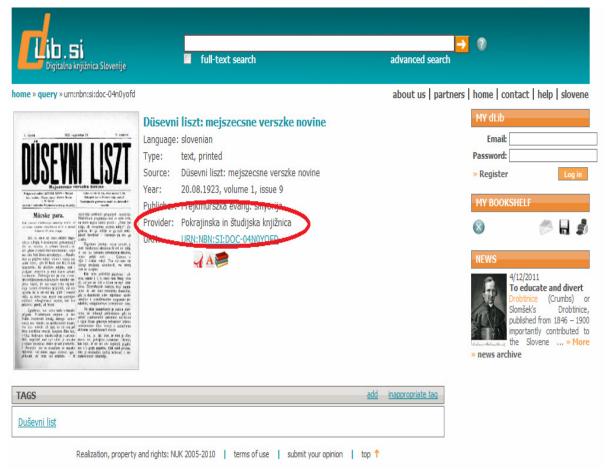


Figure 4: Example of the provider of digitised material

4. National e-content aggregator in the field of culture

In the Strategy of development of information society in the Republic of Slovenia - si2010 the Government in 2007 stated the vision of e-contents – to provide e-content in the Slovenian language in all areas that significantly affect the life and work of an individual. It also identified the objectives to realize the vision shaped the development of: providing a sufficient range and quality of e-content in the Slovenian language, the establishment of service-oriented infrastructure, enforcement of the standards of website design tailored to the users needs, provide a clear policy in the enforcement of copyright protection for e-content and enforce privacy of all web portals and providing citizens the widest possible access to adequate levels of quality e-content in foreign languages as a fundamental material for scientific, research and other cultural activities. For the cultural sector was particularly vulnerable, it was important to accelerate digitisation of existing analogue content, integration into the European digital cultural space and further development of the Digital Library of Slovenia.

Quite to the newly created European integration (European digital library) and created a more in-depth cooperation among the key players in the field of culture and science (libraries, museums, galleries, audio and video archives, universities, institutes, ...) and the desire to coordinate creating e-content (digitised and born digital) produced with public funds are re-

flected in the decision of the Ministry of Culture to establish a national e-content aggregator in the field of culture and assigned the task to the National and University Library¹.

European digital library defines aggregator as "any organization that collects metadata from a particular group of providers of cultural content and distributes it to Europeana. Aggregators collect materials from various organizations, the collected files and metadata, standardized, and, in a form that is in accordance with the guidelines and procedures of Europeana are transmitted. Aggregators also support the providers of cultural content through the administration, operation processes and training."

In the frame of the project to establish national e-content aggregators in the field of culture the analysis of the status quo of storage and distribution of e-content³ has been conceived and carried out as the basis on which it is possible to sharpen the scope of the tasks of national aggregators. It has been found that in the cultural sector in Slovenia, currently there is no umbrella organization or institution or body that would provide appropriate guidance for cultural public institutions to create strategies to help their own digitisation of cultural content. It has been estimated that there is little education, counselling and outreach content digitisation the extent and why it is important to implement digitisation. Certainly, more effort will have to be put into education and promotion of digitisation. The analysis has shown that in the Slovenian online space there are some major web portals with content from the cultural sector and the potential to become the central, particular field portals for specific cultural events and from which a national aggregator can draw the contents of national importance for the transmission to Europeana. Many public institutions in the field of culture were involved in international projects, but the analysis found the lack of coordination, cooperation and communication of results among the various digitisation projects working in related or similar fields and as a result provide related e-content that are sometimes even competitive. All the above indicates the need to develop a national strategy to create digital content in the field of culture and heritage, by which it could be followed by participating institutions. Creating an additional obstacle to faster e-content is copyright regime. It would be necessary to update and adapt the copyright law to allow for quick and easy obtaining of licenses to carry out the digitisation of copyrighted materials for the online access, permanent retention, protection and preservation of cultural heritage.

4.1. Project starting points. The above-mentioned definition and functions of aggregators indicate some basic activities that should be implemented in the context of Europeana aggregator, but certainly are not limited to activities that aggregator, or institution that manages it can do to achieve additional benefits and objectives of the country itself. Therefore the design of the project to establish a national aggregators in Slovenia defined its purpose as broader than in the context of Europeana, to "provide a single point of online access to Slovenian econtent in the field of culture, and begin by establishing the conditions for permanent storage of e-content in the area of culture and science." There were well defined project objectives, namely:

- provide a single, continuous access to e-content in the field of culture and science,
- provide open and easy access to the Slovenian e-content in culture via the World Wide Web,

¹ Ministry of Culture Decision, no. 6120-223/2008/32, 14.9.2009.

² Europeana Aggregators' Handbook. Edition 1 May 2010. http://group.europeana.eu/web/guest/provide_content/

³ NUK, IPMIT: Analiza obstoječega stanja hrambe in distribucije e-vsebin, interno gradivo NUK, 2010.

- stimulate the creation of e-content in public institutions in culture, particularly those that until now have not had technology, personnel and other potential entry into the digital environment,
- establish a connection and synergy (co)operation between the existing and future portals that provide access to e-content in the field of culture and science,
- facilitate the transfer of Slovenian cultural content in digital form (e.g. digitisation of 3D objects, digitisation of cultural content in the field of musical art, artistic creation, performing arts, etc.),
- provide a platform for creating conditions for permanent storage of Slovenian econtent in the field of culture,
- develop modalities for the provision of digitised works and facilities available to the public,
- ensure the functioning of national aggregators as links to the portal Europeana.

The emergence of digital content and services has changed the way of life, needs and aspirations of society in general. Classic services should be changed in the direction of desires and needs of users, in order to maintain their value in new situations. Development in information and communication technologies (ICT) field requires the public sector to implement change or introduce innovations in all work processes. This is also applied to institutions in the field of cultural heritage, faced with the need to introduce changes which will meet the needs of many of their users and also attract new users. Digitisation also provides new developmental challenges as opportunities for cultural institutions in fulfilling their mission in society. It is clear that cultural policy is considered to computerization, it is the creation of digital cultural content and the development of e-culture, as one of the most important overall priorities, which permeates the entire cultural sphere. It should be understood as the integration of ICT into basic processes of creation, preservation, dissemination, accessibility and reuse of digital cultural content in both direct creators of cultural goods as well as in libraries, museums, galleries, archives, media and other cultural institutions.

Digitisation, digital content, digital services and public access to culture, media and cultural heritage are not only crucial for the marketing of these and related areas of creative industries and tourism economy, but they also represent a basis for the direct use of digital cultural content in education, training, lifelong learning and research. Digital cultural content is one of the most effective instruments for raising the visibility and competitiveness of the country and its economy as a whole. In this context, cultural industries have not negative effects, on the contrary, they can support creativity. Public interest in e-culture involves the development and wide availability of digital cultural content as a basis for the recognition of Slovenia in the world, to create an additional value in the field of culture and related cultural industries and contribute to the creativity for a better quality of life.

Establishing national aggregators is therefore an opportunity to establish in the field of culture a basis for a system that will enable the achievement of development goals of econtent and e-culture. Currently, in Slovenia to offer e-content and e-services in the field of culture failure (despite the relatively large range of portals and web sites). Institutions in the field of cultural heritage, and often in this area operate piecemeal, uncoordinated, with no clear objectives and development vision. The digitisation process at the national level is inadequately co-ordinated, leading to increased costs, time losses in the digitisation of sensitive materials, less publicly available material in electronic form, the tower from knowledge of a lower quality of digitised materials, etc. The mismatch is evident in the area of metadata, where differences exist in the standards between different cultural areas (museums, libraries,

archives, etc.) as well as within areas themselves, and lack of coordination with the institutions at the EU level is even higher. In addition, the fragmented storage capacity for e-content, a lot of different platforms with diverse content and functionalities, leading to confusion and time losses to consumers, the situation in respect to copyright and privacy compliance is transparent, etc.. Implementation of national aggregators, as conceived and proposed, will ensure a rapid progress in the following areas of e-culture:

- encourage the digitisation of all types of cultural content,
- appropriate documentation,
- development of Slovenian terminology and
- permanent storage of digital content;

And in particular:

- the establishment of Internet access in one place for content from the all cultural sectors in Slovenia, which will be connected with Europeana and other stored digital cultural content in the world and
- training and development of e-skills in different groups of users in Slovenia.

It should be noted that the cultural institutions in Slovenia are mainly public non-profit organizations whose mission is to strengthen cultural identity, education and leisure time of society, so profit is not a prerequisite for defining objectives. However, public institutions must also behave economically, so the rational use of resources is also one of the conditions in the search for optimal solutions to achieve those goals. The use of any model that would bring revenue is not ruled out if it would not affect the achievement of the basic objectives of the project.

- **4.2. Aggregator organizational model.** To achieve the objectives set out in the previous chapter, we need an independent body, named all the Digital Cultural Heritage Centre (DCHC) to implement several processes. In carrying out the process the Centre appears in various roles, with the following identified key roles and processes within which the institution will be implemented.
- **4.2.1 Cultural content digitisation coordination on the national level.** Digitisation means the first step of providing access to digital forms of culture as a process of digitisation of econtent created. The current state on the digitisation of cultural content is completely decentralized, where each institution is responsible for digitising their own content, that is, the material they store and allow physical access to them. There is not a lot of cooperation, coordination and knowledge transfer between different institutions. It might therefore be appropriate that the scope of digitisation of cultural content is more centralized, and could play a central role also referred to an independent organization.

The coordinating roles of digitisation on a national level, such an umbrella organization, implement the following processes:

- to carry calls for the digitisation of Slovenian cultural heritage,
- development of standards of digitisation of Slovenian cultural heritage,
- training and consulting in the field of digitisation and the creation and use of e-content for the institutions of culture,
- training and consulting in the field of copyright law in connection with the digitisation and making available to the public.
- **4.2.2** *E-content aggregator in the field of culture.* The role of aggregators refers to the collection and distribution of metadata on the cultural content in electronic form as well as coop-

eration with Europeana. The role of national e-content aggregator in the field of culture if the organization would perform the following topics:

- the standardization of metadata about e-content in culture,
- automatic aggregation of metadata about e-content in culture,
- transmission of metadata about e-content in culture to Europeana,
- creation of a legal model for the use and transfer of metadata between institutions in the field of culture and national aggregators,
- participation in the management of Europeana.
- **4.2.3** Long term preservation of digital cultural heritage. The role of permanent preservation of digital cultural heritage of the Centre implemented the following topics:
 - manage the national repository of digitised cultural heritage,
 - the collection of the so-called mandatory digital copies (copies of files created in the process of digitising the purpose of permanent preservation).
- **4.2.4 Service provider.** In connection with an application service provider e-content in the cultural centre would perform the following topping:
 - management of the public portal Slovenian digital cultural heritage,
 - enable roaming website Slovenian institutions in the fields of culture,
 - transmission of high-quality archival versions of digital cultural heritage on the request.
- **4.2.5** Research and development. One of the objectives of DCHC may also concern the continuing progress in the fields with which to deal. Thus the Centre will carry out research that will further improve its solutions and other processes to be carried out to reduce the burden on content providers and users and further increasing involvement in the process of creating e-content and make them available.

Although from the above this is a great concept, one should not overlook any threats. The major one may include providing sustainable human and financial resources for the operation and maintenance, limited access to resources or knowledge. Content due to the rigidity of copyright law (specific content only available in the deposit unit), risk of aggregation of aggregators (Europeana eventually become the sole provider of Slovenian cultural heritage), dominance of competing private equity products (Google Books), etc. On the other hand, it is possible that the impact of economic crisis will act as an opportunity and stimulus for reflection, which will lead to an alternative solution, a solution based on integration of partial solutions and building a single, integrated system under one title – in our case, maintaining econtent in the field of cultural heritage.

References

- [1] Smernice za digitalizacijo knjižničnega gradiva [accessed 2011-04-17]. Avaliable: http://www.mk.gov.si/fileadmin/mk.gov.si/pageuploads/Ministrstvo/Drugo/aktualno/2010/smernice_za_digitalizacijo_konena_TL_ZK_MS.pdf.
- [2] Program dela 2010 [accessed 2011-04-17]. Avaliable: http://www.nuk.uni-lj.si/dokumenti/2010/pdf/NUK Plan 2010.pdf.
- [3] Razpisna dokumentacija: javni razpis za izbiro usposobljenih kandidatov za digitalizacijo tekstovnega knjižničnega gradiva. [accessed 2011-04-17]. Avaliable: http://www.nuk.uni-lj.si/dokumenti/2008/digitalizacija 1 faza-razpisna dokumentacija.pdf.
- [4] Okvirni sporazum [accessed 2011-04-17]. Avaliable: http://www.nuk.uni-lj.si/dokumenti/2008/digitalizacija_okvirni_sporazum.pdf.

- [5] Digitalna knjižnica Slovenije dlib.si [accessed: 2011-04-17]. Avaliable: http://www.dlib.si/v2/Default.aspx.
- [6] Europeana Aggregators' Handbook. Edition 1 May 2010 [accessed 2012-01-16]. Avaliable http://group.europeana.eu/web/guest/provide_content/.
- [7] Strategija razvoja informacijske družbe v republiki Sloveniji si2010 [accessed 2012-01-16]. Available: http://www.mvzt.gov.si/fileadmin/mvzt.gov.si/pageuploads/pdf/informacijska druzba/si2010.pdf.
- [8] (2006) "Zakon o obveznem izvodu publikacij (ZOIPub)", Ur. list RS št. 69/06 ZOIPub [accessed 2012-01-16]. Avaliable: http://zakonodaja.gov.si/rpsi/r06/predpis ZAKO3606.html.
- [9] (2002) Consultative, Committee for Space Data Systems, "Reference Model for an Open Archival Information System (OAIS)" [accessed 2012-01-12]. Avaliable: http://public.ccsds.org/publications/archive/650x0b1.PDF.
- [10] NUK, IPMIT: Analiza zahtev za vzpostavitev Informacijskega sistema digitalne kulturne dediščine, interno gradivo NUK, 2010.
- [11] NUK, IPMIT: Organizacijski in poslovni model nacionalnega agregatorja, interno gradivo NUK, 2010.
- [12] NUK, IPMIT: Analiza obstoječega stanja hrambe in distribucije e-vsebin, interno gradivo NUK, 2010.

<u>zoran.krstulovic@nuk.uni-lj.si</u> <u>mojca.savnik@nuk.uni-lj.si</u>