Alenka Kavčič-Čolić

National and University Library, Ljubljana, Slovenia

APPROACHING DIGITISATION THROUGH A DIGITAL PRESERVATION PERSPECTIVE

Abstract: The European Commission has been supporting the digitisation of cultural heritage for more than 10 years. As result, Europeana, the joint portal of European libraries, has already overpassed 20 million digital objects. Many funds have been invested in different national and international digitisation projects; their main concern was to provide public access to different types of content. Most libraries still conceive digitisation as a digital reproduction aimed to provide access to library materials only. The master files resulted from digitisation are usually not digitally preserved and the digital collections run the risk of being lost for the future. This is even more worrying since digitisation, especially of old library materials, is a very expensive task and it is very probably that we will not be able to repeat the scanning process for some time. For this reason, the produced digital collections should be treated as an important asset to the library.

In this paper, the importance of preserving the digitisation master files or, using Conwell's terminology, to "digitise for preservation", by introducing digital preservation methods usually applied on digitally born materials will be discussed. Arising from the digitisation experiences of the National and University Library in the European projects EOD and IMPACT, we had discussed the need to change the approach to digitisation and start producing digital master files that would satisfy the requirements for digital preservation. We will focus on the most decisive elements that have influence on extending digital master files life cycle, i.e. digitisation project planning, digitisation tools, quality of the image and metadata needed to preserve the master files.

Keywords: digitisation, digital preservation, digitisation planning, digitisation projects

1. Introduction

Digitisation has become an intrinsic task for modern libraries; one of their priorities is to build digital collections and provide online access to their contents. During the last ten years, the European Commission (EC) has been supporting digitisation on policy and strategic levels [6] (Lund Principles, Lund Action Plan in 2001,¹ the Dynamic Action Plan 2006-2010², i2010³, The Digital Agenda for Europe 2010-2020⁴). The EC has also had a very important role in coordinating and funding digitisation projects in the European Union countries through different development and research programmes (eContent⁵, eContentPlus⁶, CIP-ICT-PSP⁷ and ICT 7th Framework Programme⁸). The biggest European joint project that has had major impact on digitisation in Europe was the creation of the European Digital Library – Europeana⁹ which provides access to more than 20 million of digital objects from every European country.

All these actions and initiatives have contributed to the development of the information society and to conceive digitisation as a usual way of reproduction of analogue library materials. Digitisation in libraries has been oriented mainly to production (quantity) and less to digitisation quality assurance.

¹ Available at URL: http://cordis.europa.eu/ist/digicult/lund-principles.htm

² Dynamic Action Plan for the EU co-ordination of digitisation of cultural and scientific content, available at URL:

http://ec.europa.eu/information_society/newsroom/cf/itemdetail.cfm?item_id=2358&utm_campaign=isp&utm_medium=rss&utm_source=newsroom&utm_content=tpa-115

³ Available at URL: http://ec.europa.eu/information_society/eeurope/i2010/index_en.htm

⁴ Digital Agenda for Europe (COM(2010) 245 final/2): http://eurlex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2010:0245:FIN:EN:PDF; website:

http://ec.europa.eu/information_society/digital-agenda/index_en.htm

⁵ http://cordis.europa.eu/econtent/

⁶ http://ec.europa.eu/information_society/activities/econtentplus/closedcalls/econtentplus/index_en.htm

⁷ http://ec.europa.eu/information_society/activities/econtentplus/index_en.htm

⁸ http://cordis.europa.eu/fp7/ict/programme/challenge4_en.html

⁹ http://www.europeana.eu/portal/

In 2002, UNESCO published the first guidelines for digitisation [12] which had no major impact on libraries (Quality assurance..., 2007, p.40). A few years later the first standards, best practices and guidelines were adopted; however, the digital conversion quality differed from one library to another. Hurst-Wahl [15] calls the first years of digital conversion as the "Wild West" days of digitisation. During the last few years the situation has somewhat changed and the scanning technology has considerably improved.

A survey undertaken by the Dutch Centre of Expertise [25] on the current state of affairs regarding quality assurance in relation to the digitisation of cultural heritage in twelve European Union member states, revealed that although some of the countries ascertained the need to adopt policies and standards in the field of digitisation, no one of them has officially done anything in this field. The Minerva project¹⁰ partners (running till 2008) started some actions in fostering the adoption of international standards to assure quality of digitisation. They published a Good Practices Handbook [11] and they started a website with valuable information on digitisation.¹¹ Since then, there was a proliferation of guidelines and centres of competence aimed to disseminate best practices and guidelines related to digitisation and adopt relevant standards in this field. However, most digitisation projects were focussed on access and not on preservation quality. As consequence, the survival of the produced digital collection might be in many cases endangered. This is even more worrying since huge amount of resources have been invested in digitisation.

2. The reasons of digitisation for preservation

Preservation in the context of digitisation is very often understood as preservation of the original library material in a way that its content is accessible to the users but the physical item remains safe from degradation and damage produced by its use. Some experts [2, 4, 23, 28, 8] believe that the digitisation master files should be preserved like digitally born materials. Conway [7, 8], for instance, considers digitisation as "an investment in the creation of lasting digital products" (p.65). He introduces the term "digitisation for preservation" for denominating the process of creating "valuable new digital products and protecting the value of those products, regardless of whether the original source is a tangible artefact or data that were born and live digitally." (p.64-65). The Association for Research Libraries (ARL) »supports digitization as an acceptable preservation strategy« [2]. According to ARL, every digitisation work (»reformatting method«) is a unique procedure in which we have to make choices that are »based on the characteristics of the original, the capabilities of each reformatting process, the current and anticipated needs of the user, and costs« (p.1)

Huge investments in digitisation projects cannot be disregarded. Many libraries heavily raise funds or separate their budgets for this task. The NUMERIC survey on sources of digitisation funding revealed that 62% of libraries fund digitisation projects from their own resources and 30% form government programmes [20].

There are many endeavours in assessing the digitisation costs [23, 17, 3, 13, 5, 21, 20]. They are based on different methodologies and provide quite different figures. In their calculations, they include different categories of tasks like: digitisation planning, training of image capture and processing staff, selecting and implementing a system, developing and maintaining software, retrieval of volumes from storage, identification, collation and repair, handling of original library materials, scanning, image processing, metadata creation, OCR, SGML, or TEI encoding, quality control, collection building, and system maintenance.

The project Making of America [3] revealed that the digital conversion was the third part of all digitisation costs. The rest included metadata creation (29%) and other activities costs (39%).

A considerable reduction of digitisation costs has been achieved with robotic scanners, which can digitise a book by turning its pages automatically. These scanners are becoming more and more fast,

¹⁰ http://www.minervaeurope.org/

¹¹ See http://www.minervaeurope.org/home.htm

producing better images as well. However, even if there is a very detailed selection process, and if it is based on standards and good practice guidelines, the primary priority of large-scale digitisation is to create a critical mass of digitised books, often on the expense of the highest quality.

A few years ago, two major commercial companies, Google and Microsoft, started the biggest digitisation projects Google Book Search program¹² and Microsoft Live Search Books¹³. Both had an important impact on the library community and have contributed to providing access to valuable digital collections. But as large-scale digitisation projects, the quality control was done on a sample of scans. According to Helm [14], in Google project "*librarians and academics say some early scans include a variety of errors, such as blurred words, missing pages, and truncated text in public-domain texts from the library program. (...) Digital versions of books scanned using the Google system contains far more errors than pages scanned loose-leaf.*" Hurst-Wahl [16] was also critical regarding Google's quality at the beginning of its project.

Digital preservation is not cheap and we do not have a perfect solution yet – it is changing all the time. According to Walters and Skinner [29] "there likely will never be a perfect solution in technical terms; more likely we will develop a set of evolving curatorial processes that help ensure the usefulness of these digital objects for present and future generations" (p.260). Therefore, the question is not how much it will cost us if we do it, but how much it will cost us if we do not do it. If libraries do not take a long-term care of these resources they run the risk of losing money, time and effort that went into the creation of them, as well as losing the present and future researchers who depend on these resources. "They run the risk of undermining their own stability as institutions" (p.262)

Many librarians entirely rely on commercial companies to which they outsource all the digitisation work. It may happen that these commercial companies do not share librarians' values and commitment to preserve digitised material [4]. That is the reason why librarians have to take active role in digitisation projects and require the production of highest quality digitisation master files which will help to preserve the digital collections for the future.

In the European project IMPACT¹⁴ (run till 2011) we needed a set of digital master files of old digitised books in file format TIFF or at least in JPEG 2000 with a lossless compression, accompanied by a 100% accuracy of optical character recognition (OCR). A part of the budget was designed for outsourcing this task. After we sent the request with our desired specifications to several commercial companies, some of them tried to convince us that for access we do not need TIFF or JPEG 2000 quality, but it is enough JPEG with loss compression, which is what they usually do for libraries, and their OCR which is 95% is sufficient for full-text searches. They refused to pack the metadata in the specified XML schema, because it was too time consuming. At the end, we decided to do all these required tasks by our own. It was an excellent experience, because we learned a lot about the tasks, we understood the different level of processes in each task, and in the future it was much easier to negotiate with the companies. This is an example showing why librarians need to take an active role in this field by taking charge and control of the whole digitisation work. And further, they need to require the best quality of the master files, since the investment in digitisation is very high. It is better to get less digital copies with the highest quality that can be used and reused in the future.

3. Preservation approach to digitisation

In 2005, the National and University Library (NUK)¹⁵ started building its Digital Library. In May 2012 NUK total online digital collection comprised more than 3,5 million scans, including: 19.000 pages of scientific journals, 402.714 pages of newspapers, 9.540 photographs, 100 music records, 15 3D objects, 3 virtual exhibitions etc. NUK is an active partner in Europeana and other international projects in the field of digitisation, like eBooks on Demands (EOD)¹⁶ and IMPACT. It combines

¹² http://www.google.com/googlebooks/about.html

¹³ http://books.live.com

¹⁴ http://www.impact-project.eu/

¹⁵ http://www.nuk.uni-lj.si

¹⁶ http://books2ebooks.eu/sl

outsourcing and in-house digitisation, which is the source of know-how and rich experiences in digitisation. As the main Legal Deposit institution in Slovenia, NUK has recently adopted a Digital preservation strategy and maintains a digital archive for library materials. In continuation, the experiences that resulted from our direct involvement in the EOD and IMPACT projects will be presented.

The digitisation master files of NUK's digital collections are important assets of the library which maintains their life cycle in the same way as of the digitally born materials. We only regret that when we started with digitisation projects we did not give such attention to preservation requirements. As consequence, we have master files of earlier digital collections in non-archival formats, which might cause problems in the future.

Digitisation for preservation can slow the digitisation project and requires more investments during and after the project. Because of that, it is very important that the digitisation project is supported by the management and the whole organisation. The head of the digitisation project should not be only a good manager, but should very well understand every digitisation process and support digitisation for preservation concept.

Digital preservation starts from the very moment when the digital object is produced otherwise we can lose not only the quality but also all the valuable information that is related to the content and is important for its preservation. According to Rieger [28], "*digital preservation requires a sequence of decisions and actions that begin early in an information object's life cycle.*" (p.15). For this reason the planning stage is the most critical in a digitisation project. Usually the digitisation projects are structured in the next stages:

- 1. Materials selection and acquisition
- 2. Metadata definition
- 3. Image capturing
- 4. Image processing
- 5. Additional metadata generation and encoding (OCR, TEI ...)
- 6. Quality evaluation
- 7. Digital collection management

In the planning stage of the digitisation project, we need to clarify:

- what are the aims, objectives and target audiences of the digitisation;
- what are the characteristics of the material we need to digitise;
- has the material already been digitised somewhere else if yes, what is its quality;
- do we need with intellectual property rights clearance;
- what kind of software and scanning tools we need for getting the best quality of digitisation;
- what kind of metadata will be used for describing the items;
- what will look like the final digital collection; what kind of functionalities it will have;
- will it completely replace the original library material;
- does it need additional conservation work;
- and, finally, how many resources we need; and
- how much it will cost.

It is not in the scope of this paper to describe in detail the digital preservation methods or present the digital preservation workflows in a library that has mandate to preserve its digital resources. However, some of the most decisive elements that influence the quality of digitisation and, therefore, the long-term durability of the digital object will be addressed. These are: scanning technology, image quality and metadata specification.¹⁷

¹⁷ For quality digitisation projects we recommend some useful guidelines and handbooks [1, 9, 10, 12, 18, 19, 30].

3.1 Scanning technology

Scanning (image capturing or digital conversion) can be carried out in-house in the library premises or outside, by an outsourcing company, depending on the type, characteristics, and quantity of the material to be digitised. The quality of digital image mostly depends on the scanning technology, scanning software, and trained operational personnel. Before outsourcing scanning, it is recommended to get enough information on the experiences of the outsourcing company, their staff and technology. It is also very important to specify in detail all the requirements.

If it has been decided to buy a scanner for in-house digitisation, there are some requirements that the scanning technology (scanner or digital camera) should satisfy:

- it needs to achieve at least 24 bit colour and 600 dpi resolution;
- it should support main archival and lossless file formats (TIFF 4 or PNG for pictures and JPEG 2000 for texts) for the master files to be preserved;
- it can accommodate the physical dimensions of the originals (different dimension formats);
- it can support the physical condition of the documents (cold light, page turner, scanning speed, etc.) without damaging them;
- it supports different image processing software.



Picture 1: Scanned image before processing....



Picture 2: Scanned image after processing

Pictures 1 and 2 show differences between pre- and post-processing.

In addition to the scanner, good digitisation software for capturing, processing and delivery of digital images is needed. Usually, a basic version is already integrated in the scanners and many companies offer additional upgrades, which are almost the same price as the scanner. It is very important that we have the necessary scanning tools for achieving the highest quality of the image. During the EOD project we tested different image processing software and we realised that they differed not only in the functionalities, but also in the colour treatment.

3.2 Image quality

The resolution and bit depth contribute to the image quality. Every image is compound of picture elements or pixels (a single point or tiny square in a graphic image stored in an ordered rectangular grid – the smallest element in a digital image). Resolution is determined by the number of pixels used to present the image, expressed in dots per inch (dpi) or pixels per inch (ppi). We need to distinguish *input* from *output* resolution [30, 18]. The input resolution is the one that is produced by the scanner or digital camera, while the output resolution depends on the capabilities of the computer screen or printer. In digitisation for preservation it is important to get the maximal input resolution and produce archival file formats. For display and use, a secondary access file format will be produced depending on the output resolutions and user requirements.

When defining the input resolution we need to know the dimension of the material we want to scan. If for instance, we scan an image with the resolution of 300 dpi, and for accessing to it we need to enlarge it twice, its output resolution will be 150 dpi. To obtain the same quality, we should digitise it at a 600 dpi of resolution and the way round (Picture 3).



Picture 3: Example of different digitisation resolutions

Another issue that influences on the quality of digital images is the bit depth. Zhang & Gourley [30] define bit depth as the "(...) amount of information (stored as bits) that is represented by a single pixel." The digital colours are represented by a binary value (0, 1). If we use 1 bit depth, we will get a bitonal image (black and white). With a higher bit depth we can produce grayscale or colour images. The difference is whether the intermediate tones represent the black to white scale or of red, green and blue (RGB) scale. A colour bit depth of 24 means that the colour images have 8 bits per colour (2^8) or 256 tones per each colour, making together 16.777.216 (224) of different colours (See Picture 4 and Table 1). This is called »true colour« because it is in the range of the perception of human eyes. It is also the highest depth available on normal computer screens.



Picture 4: 24 bits colour depth

For archival purposes when digitising black and white texts it is recommended to use 8 bits grayscale and for coloured pictures a 32 to 48-bit colour scale in order to preserve all the details.

Bit	Number	Example of digital values of each colour	
depth	of colours avaliable		
1	2	1, 0	Black / white
2	4	11, 00, 01, 10	Grey scale
			image
4	16	000, 0001, 0010,	
8	256	0000000, 00000001,	
16	65.536	00000000000000, 0000000000000001,	
24	16.777.216	00000000000000000000000,	'true colour'
		000000000000000000000000000000000000000	
		000000000000000000000000000000000000000	
		0000000000000000000111	
48	281 trillion	000000000000000000000000000000000000000	

Table 1: Bit depth and number of colours ratio in a digital image

3.3 Metadata specification

There are different categories of metadata related to the digital object. Descriptive metadata is used for identification, access, discovery, and management of digital resources. Usually, it is the bibliographic record in the catalogue. Administrative metadata includes rights and permissions. Technical metadata refer to the characteristics of the digital object. Structural metadata provides information about the internal structure of resources and relationships of the integrating parts of complex digital objects and their relationship with other digital objects. For digital preservation, specific metadata are required which include some of the above mentioned categories. These metadata are based on the information model of the Open Archival Information System (OAIS) reference model (ISO 14721/2003) [26]. The key preservation metadata schema is The Data Dictionary for Preservation Metadata (PREMIS -PREservation Metadata: Implementation Strategies). It defines preservation metadata as data that »supports the viability, renderability, understandability, authenticity, and identity of digital objects in a preservation context; « which »represents the information most preservation repositories need to know to preserve digital materials over the long-term;« [21]. Most of the technical metadata important for preservation can be automatically extracted from the digitisation master files, with the exception of some specifications regarding the image production, like used scanning technology, image processing tools or documentation on the interventions and decisions regarding the digitisation selection, scanning, selection of file formats etc. According to Rieger [28], "documenting the attributes of digitized materials in a consistent way makes it possible to identify the provenance of an item as well as the terms and conditions that govern its distribution and use." (p.20-21).

This information can be important for quality assessment and error repair. The preservation metadata of similar items in a collection can be captured at the collection level. Unfortunately, due to giving priority to access, most libraries do not pay attention to these metadata.

In addition to the described preservation metadata, persistent identifiers should be assigned to digital objects. They are long lasting references to digital objects. There are many types of persistent identifiers that can be based on an algorithm or a code number. To verify fixity of a file, level checksum can be used.

In digitisation the optical character recognition (OCR) is very important, especially in large-scale digitisation projects, since it produces the text that is the basis of full-text searches. OCR is defined as "A process by which characters typed or printed on a page are electronically scanned, analysed, and if found recognizable on the basis of appearance, converted into a digital character code capable of being processed by a computer." [27]. OCR technologies cannot achieve 100% accuracy yet. For a better OCR % a perfect quality of the scanned images and additional post-correction are required

(Picture 5). Special problem are imperfections in old texts caused by stains on the paper, foxing, humidity (paper wrapping) (Picture 6), blurred, broken, faded characters, etc. The IMPACT project addressed some of these problems and developed software that improved OCR in several European old languages. It certainly requires more work on the images and OCR quality, and it should be taken into account in the planning stage of the digitisation project.

Ni vse zlato, kar se sveti! (Kako si naši čeitatčelfii ramzliagafjo naša gesla?) NARODNA SLOGA i - Zares bogata vsebina je v tem <u>zl'</u>0atem rekuji i≛n globoka življenska mo-+ drost je njega nazorna vsebi<u>nim</u>! • <u>D</u>Pobili sm.no več spestavkov v tem_-predrnmetu, zato prinašamo le odlomrhke, da <u>I fes</u>e mikKom-n-ur iinee zaramerimo; kajti veliki krog naših čitateljev je baš tako zaacnmerljiv, kakor ženska, če razpravljasš to njenih leti<u>h</u>li. Ker smo vsi ljii, udje pod kožo krvavi -in 'nagnjeni po jnaravi k raz.iiVn<u>nim</u> <u>s(5</u>1a- ibostim in i tudi krepostim n, bodo rafcftčurazlična mnenj^a naših s6odelav-ce/v pr-av -verno zrcalo mtnišijenja naših čitateljtev. Po praviluhi gl^as ljudstva je glas - bož-IGlas ljudstva je glas božji — (vox) popubli", vox deci). Po temtn pravil-u objav-Ijamo na kratkod v ifzvlečkih sledeča tnmnenj[a:

Picture 5: The image of *Narodna sloga: neodvisen tednik*, (27.02.1932, vol. 2, nr. 8), automatically processed with OCR and additionally manually corrected



Picture 6: Imperfections caused by humidity (paper wrapping)

The produced OCR text from digitised images and different categories of metadata can be encoded in different schemas METS¹⁸, TEI¹⁹.

4. Conclusion

From its very beginning, digitisation in libraries and other cultural heritage institutions has been considered a way of digital reproduction of analogue materials, and has been connected to provision of access to their holdings and building of their digital collections. After two decades, it is time to reconsider these views. Firstly, digitisation is a very expensive task and libraries hardly invest funds not only for building but also for maintaining their digital collections. Secondly, the digitisation technology is already sufficiently developed for obtaining high quality copies. Thirdly, the digital preservation community has already adopted standards and preservation workflows that can be integrated with the digitisation process. The digital preservation process (or "digitisation for preservation") starts at the moment of producing the digital image. For this reason, it is necessary to pay more attention to the planning stage of the digitisation project and set higher requirements regarding the most decisive elements in the creation of the digital image, that is: proper scanning technology which avails to produce digital images of the highest quality in archival file formats, the highest resolution and bit-depth that can be adjusted to the type of library material, and the metadata, which include all the information important for extending the digital files life cycle. Librarians can contribute a lot by taking active part in the digitisation workflows and demanding the best quality from digitisation operators. They need to start thinking about their digital collections as if they were the most important assets that increase the importance of their library to the community and stakeholders they serve.

References

[1] Abu-Zayed, Ahmed (2009). *Digitization workflow and guidelines : digitization processes*. Exeter : University of Exeter. Visited on 3 Aug. 2011 at URN:

[2] Arthur, Kathleen ... [et al.] (2004). *Recognizing digitization as a preservation reformatting method*. ARL. Visited on 30 Apr. 2012 at URN: http://www.arl.org/bm~doc/digi_preserv.pdf

[3] Assessing the costs of conversion : Making of America IV : The American voice 1850-1876. A Handbook created for the Andrew W. Mellon Foundation (2001). Michigan : The University of Michigan Digital Library Services. Visited on 30 Apr. 2012 at URN: http://www.lib.umich.edu/files/services/dlps/moa4_costs.pdf

[4] Bellardo Hahn, Trudi (2008). Mass Digitization. Implications for Preserving the Scholarly Record. *Library Resources and Technical Services*, 52(1), pp.18-26.

[5] Bia, Alejandro, Muñoz, Rafael, Gómez, Jaime (2010). DiCoMo: the digitization cost model. *International Journal on Digital Libraries*, nr. 11, pp. 141-153.

[6] Commission Recommendation of 24 August 2006 on the digitisation and online accessibility of cultural material and digital preservation (2006/585/EC). *Official Journal of the European Union*, L 236, 31/08/2006 pp. 0028 - 0030

[7] Conway, Paul (2000). Overview: Rationale for digitization and preservation. In: Maxine K. Sitts (Ed.) *Handbook for Digital Conversion Projects: A Management Tool for Preservation and Access*. Andover, MA: Northeast Document conservation center. Visited on 30 Apr. 2012 at URL:

http://www.nedcc.org/resources/digitalhandbook/ii.htm

[8] Conway, Paul (2010). Preservation in the age of Google: digitization, digital preservation, and dilemmas. *Library Quarterly*, 80(1), pp.61-79.

[9] *Digitisation Guidelines* (2004 -). The National Library of Australia. Visited on 4 Aug. 2011 at URL: http://www.nla.gov.au/standards/digitisation-guidelines

[10] DFG Practical Guidelines on Digitisation for programmes funding Scientific Library Services and Information Systems (2009). / The Subcommittee on Cultural Heritage on Cultural Heritage. Bonn : Deutsche

http://projects.exeter.ac.uk/charter/documents/digitizationWorkflowGuidev5.pdf

¹⁸ Metadata Encoding and Transmission Standard: http://www.loc.gov/standards/mets/

¹⁹ Text Encoding Initiative: http://www.tei-c.org/index.xml

Forschungsgemeinschaft. Visited on 3 Aug. 2011 at URN:

http://www.dfg.de/download/pdf/foerderung/programme/lis/praxisregeln_digitalisierung_en.pdf

[11] Drake, Karl-Magnus, Justrell, Borje, Tammaro, Anna Maria (2003). *Good Practices Handbook*. Ver. 1.2. / edited by the Minerva Working Group 6: Identification of good practices and competence centres. Visited on 30 Apr. 2012 at URL:

http://www.minervaeurope.org/structure/workinggroups/goodpract/document/bestpracticehandbook1_2.pdf [12] *Guidelines for digitization projects for collections and holdings in the public domain, particularly those held by libraries and archives* (2002) / IFLA & ICA. – Paris : UNESCO. Visited on 3 Aug. 2011 at URL: http://www.ifla.org/en/publications/guidelines-for-digitization-projects-for-collections-and-holdings-in-the-public-domain

[13] Hammond, Max, Davies, Claire (2009). *Understanding the costs of digitisation. Detailed report*. Surrey : Curtis+Cartwright Consulting Ltd. Visited on 30 Apr. 2012 at URL:

http://www.jisc.ac.uk/media/documents/programmes/digitisation/digitisation-costs-full.pdf

[14] Helm, Burt (2005). Google's great work in progress. *Bloomberg Businessweek*, December 22, 2005. Visited on 30 Apr. 2012 at URL: http://www.businessweek.com/technology/content/dec2005/tc20051222_636880.htm

[15] Hurst-Wahl, Jill (2009). Digitization: How many best practices, guideline, and standards do we need? *Information Standards Quarterly*, 21 (4,Fall), pp.22-24. Visited on 30 Apr. 2012 at URL: http://www.hurstassociates.com/pdf/jhw isqv21no4.pdf

[16] Hurst-Wahl, Jill (2005). How is Google's digitization quality? *Digitization 101*, 30 Nov. 2005. Visited on 30 Apr. 2012 at URL: http://hurstassociates.blogspot.com/2005/11/how-is-googles-digitization-quality.html

[17] Lee, Stuart D.: Digitisation (2001). Is it worth it? *Computer in Libraries*, 21(5). Visited on 30 Apr. 2012 at URL: http://www.infotoday.com/cilmag/may01/lee.htm

[18] Note, Margot (2011). Managing image collections. A practical guide. Oxford : Chandos Publishing.

[19] Beagrie, Neil, Jones, Maggie (2008). *Preservation management of digital materials: The Handbook*. Heslington : Digital Preservation Coalition, pp. 83-151. Visited on 30 Apr. 2012 at URL:

http://www.dpconline.org/advice/preservationhandbook

[20] Poll, Roswitha (2010). Digitisation in European Libraries: Results of the NUMERIC Project. *Liber Quarterly* 19(3/4), pp.248-258.

[21] Poole, Nick (2010). The cost of digitising Europe's cultural heritage. A report for the Comité des Sages of the European Commission. Collection Trust. Visited on 30 Apr. 2012 at URL:

http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/annexes/digiti_report.pdf [22] *PREMIS data dictionary for preservation metadata* (2008). Version 2.0. Washington, DC: Library of

Congress. Visited on 30 Apr. 2012 at URL: http://www.loc.gov/standards/premis/

[23] Puglia, Steven; Reed, Jeffrey; and Rhodes, Erin (2004). *Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images.* Washington, DC: NARA ; Digital Library Federation. Visited on 30 Apr. 2012 at URL:

http://www.archives.gov/preservation/technical/guidelines.pdf

[24] Puglia, Steven (1999). The costs of digital imaging projects. *RLG DigiNews* 3(5). Visited on 30 Apr. 2012 at URL: http://chnm.gmu.edu/digitalhistory/links/cached/chapter3/link3.10b.digitalimagingcosts.html

[25] *Quality assurance in digitisation of Cultural Heritage in Europe. Final Report.* (2007). The Hague : Digital Erfgoed Nederland (DEN). Visited on 30 Apr. 2012 at URL:

http://www.den.nl/getasset.aspx?id=Website/Digitisation_of_cultural_heritage_in_Europe%5B1%5D.pdf&assett ype=attachments

[26] Reference Model of an Open Archival Information System (OAIS) : Recommendation for space data system standards : CCSDS 650.0-B-1.Blue Book (2002). / Committee on Intellectual Property Rights and the Emerging Information Infrastructure, the Computer Science & Telecommunications Board, and the Commission on Physical Sciences, Mathematics, and Applications. Washington: CCSDS (Consultative Committee for Space Data Systems). Visited on 30 Apr. 2012 at URL: http://public.ccsds.org/publications/archive/650x0b1.pdf

[27] Reitz, J.M. (2004). *Dictionary for Library and Information Science*. Wesport, Connecticut ; London : Libraries Unlimited.

[28] Rieger, Oya Y. (2008). *Preservation in the Age of Large-Scale Digitization: A White Paper*. Washington: Council on Library and Information Resources, 2008. Visited on 30 Apr. 2012 at URL: http://www.clir.org/pubs/abstract/reports/pub141

[29] Walters, Tyler O., Skinner, Katherine (2010). Economics, sustainability, and the cooperative model in digital preservation. *Library Hi Tech*, 28(2), pp. 259-272.

[30] Zhang, Allison B. and Gourley, Don (2009). *Creating digital collections. A practical guide*. Oxford : Chandos Publishing.

alenka.kavcic@nuk.uni-lj.si