eIFL Case Studies on Low Cost Digitisation Projects

Final Report

Prepared by
Repke de Vries

Edited by
Arnold Hirshon, NELINET

Table of Contents

Introduction, Methodology and Acknowledgements

This report summarizes the experiences with digitisation by some eIFL countries. Although there are probably many more examples of digitisation in eIFL countries, this report includes only those where the country responded to the survey.

The main objective of this study was to raise awareness about best practice digitisation projects that are: (1) affordable, (2) easily managed at the technical and organisation level, (3) sustainable, and (4) enable eIFL countries to preserve and promote their local content online.

Libraries in eIFL countries with digitisation projects were asked to complete a survey that asked them about the intent of their projects. Surveys were completed and returned by libraries with additional relevant information, such as pictures illustrating the digitisation (scanning). Subsequently, the respondents were interviewed briefly by phone or Skype about their survey answers.

The questions in the survey were selected from existing digitisation study questionnaires. We thank the Humanities Advanced Technology and Information Institute (HATII), University of Glasgow and Ann Gow for the use of the survey instruments they developed for the 2001 NINCH study from which a number of the eIFL survey questions were derived. The section on cost issues was inspired by the NINCH Good Practice Guide [NINCH], the iMARK Information Management Resource Kit Module on "Digitization and Digital Libraries" [iMARK] and "From paper to collection" (Loots et al., 2004)

Thanks also to Patricia Liebetrau, Geoffrey Salanje, Tigran Zargaryan for their early feedback, their pre-testing of the survey and to Jan Andrzej Nikisch who had a very thorough look at my draft texts and made valuable suggestions. Thanks also to everyone from Low Cost Digitisation Projects who answered the questionnaire and sometimes had further conversations with the author, and whose enthusiasm and detailed information, especially under sometimes tight deadlines, is greatly appreciated.

Repke de Vries
October 2008 – February 2009

Executive Summary

Libraries in eIFL countries face the same challenges as any other library in the world: how to meet new demands for digital content and at the same time maintain responsibility for older, physical collections. Digitisation and digital libraries are invaluable to answer these challenges.

Effective digitisation projects and digital library implementations require equipment, human resources and expertise. These costs involve different types of financial investments. Outsourcing is an option but usually at a higher cost. After the digitisation is completed, the library must also develop solutions and policies for access, and steps must be taken to ensure sustainability.

The case studies included in this report on Low Cost Digitisation projects collectively seek to answer this question: how can libraries in eIFL countries manage digitisation projects given the cost and policy requirements?

There are three outcomes from this study: (1) the full project reports from each of the countries that responded to the questionnaire, which was primarily for internal use, (2) this Final Report with its highlights and conclusions, and (3) a short Frequently Asked Questions (FAQ) section (in Appendix B of this report). Questionnaire used for this study is available from eIFL.net website – www.eifl.net.

The emerging picture

Out of the 47 eIFL.net countries, 20 countries were approached for information about low cost digitisation projects, of which 13 countries responded. In addition 3 countries voluntarily responded to the initial Call. The 24 countries that were not asked either were above Low Cost Digitisation project criteria or were still in the beginning stages of their projects.

Libraries responded to the survey from late 2008 to early 2009. They come from most regions of the eIFL network: Eastern Europe, Central Asia, and Africa. The conditions under which they digitise differ not only among the regions but also compared to the United States and Western Europe. Differences relate to availability of government financing, adequateness of library budgets, maturity of outsourcing solutions, and (for in-house projects) the ease of access to digitisation expertise, scanning equipment and support.

With the exception of countries that chose outsourcing, all cases report in-house digitisation through the re-deployment and retraining of staff. The libraries acquired digitising equipment and software. With a few exceptions of digitized collections of photographs, maps, and audio files, most of the digitised collections were of handwritten or printed material, mixed materials addressing "everything issued in a country about a country," and publications related to university research and teaching (such as research papers, dissertations, theses and curriculum literature). Preservation needs, together with new forms of open electronic access, drove many of the libraries to digitised manuscripts, rare books and older print journals. The scanning of contemporary print material (especially in Africa with access limited to the university network) was driven by the need to provide easier and multiple points of access. After scanning, printed text scanning can be converted from images to full text using OCR, but only two countries reported

using this approach. One project found that OCR was not adequate to cope with the country's scripts, and so they reverted to manually keying-in by volunteers.

Most scanning equipment the countries employed was low-to-mid range and was manufactured for consumer markets and office use. The few projects to digitise rare materials used more specialised high-end scanners.

The purchase of the scanning equipment and related information technology almost never could have been financed from the libraries' budgets or national funding programs. The countries typically had to rely heavily on external funding and donations, including a diverse array of initiatives and organisations including the British Library's "Endangered Archives," UNESCO, the World Bank and the Association of African Universities.

The libraries avoided personnel costs to a large extent through a combination of in-house digitisation and part-time re-deployment of regular staff. Nevertheless, there were some full digitisation teams that reported having a part-time director, metadata specialist, technical support staff, a curator, photographer or evaluation specialist and several (2 to 3) digitisers or scanner operators. Initial training for these new duties and gaining expertise was managed in many creative ways to avoid costs.

All of the reported projects were relatively young (all were from the last five years). A number are extending the length of time the project is running as a cost saving method. A few participate in international digitisation programs. One case concerned libraries that worked together as consortium on digitisation and digital libraries. Another country reported cooperation with neighbouring countries. However, most countries report that their digitisation projects occur in relative isolation.

Any digitising is a complex series of steps involving many technical and other decisions. Guidelines like the "IFLA Digitization Guidelines" offer assistance with these issues, but almost none of the cases reported using such guidelines. Once a project is finished, sustainability of the new digital collection, together with its access, starts needing attention and a budget. Most libraries recognise these sustainability issues and plan for updating their user interface, the file formats and sometimes the metadata scheme. More than half also report developing a preservation strategy for their digitised content or having one in place.

Conclusion and recommendations

The message from the reports seems to be represented by the following quotes: "we cannot afford to wait with digitisation because some of our collections are in bad condition or because we want to open up our cultural heritage to a new and broader audience or because teaching and research need much more convenient electronic copies of publications." "Therefore we [tried] a Do-It-Your-Self approach with our own staff and … not too expensive equipment." "Maybe the results are scanned images only and don't meet some of the standards … but we need digital content and need it quickly: our users and our collections are asking for it."

Against this background this case study confirms that Low Cost Digitisation is an option that can provide good results. The major hurdles are the initial costs of scanning equipment and related information technology. Many funding schemes are already helping with this obstacle. Mediation by eIFL concerning funding programs and

digitisation projects could be of further assistance. Facilitating initial training of digitisation teams could be another form of high-level eIFL assistance.

The relative isolation in which projects are done is perhaps inherent to the do-it-yourself approach, but it is less effective. As one example in the study shows, local collaboration through a consortium makes the digitisation job easier. Setting up such collaboration could be a topic for eIFL organised workshops.

It is remarkable how some of the ideas behind these Low Cost Digitisation projects have parallels in a 2007 United States forum discussion published as "Shifting Gears – Gearing Up to Get into The Flow." The message boils down to this: "vast quantities of digitized primary materials will trump a few superbly crafted [special] collections." Fortunately, the eIFL.net countries libraries represented in this study have been and are putting this principle into practice.

Low Cost Digitisation: Sample Projects

Note. Although there are probably many more examples of digitisation in eIFL countries, this report includes only those where the country responded to the survey.

Armenia. The Fundamental Scientific Library (FSL) of the National Academy of Sciences started a two year digitisation project that will scan 15.000 pages from a single collection of early printed books and periodicals that includes 400 rare books and 1.500 journal issues. The project was supported by a 47.000 GBP grant from the British Library's "Endangered Archives Programme." The one-time grant enabled the purchase of high end scanning equipment, a graphical workstation and 2TB of disk space, and training of the team's metadata specialist and digitisers by external consultants. Access to the digitised collection will be open and available through a Greenstone digital library. The project will end in 2010, and future digitisation plans are already being made to create a separate budget line for digitisation in the Fundamental Scientific Library's annual budget.

Bulgaria. The University, the City of Sofia, and the public libraries joined in a two-year digitisation effort that was fully funded by UNESCO and completed in 2006. The project resulted in significant parts of their serials collections digitised through outsourcing by a Bulgarian company.

Ethiopia. The Addis Ababa University Computer Center undertook an ambitious program in 2008 to digitise all Ethiopian collection materials that are centrally held by the university. Working with a mid-range scanner and OCR software donated by the Association of African Universities, the program is being funded through the regular budget and the worked is performed by regular staff. The use of OCR results in fully searchable text that is available and displayed next to scanned page images.

Georgia. The work of Georgian classical literature and of modern Georgian writers is being digitised under a project begun ten years ago by as a National Parliamentary Library program. The digitisation of Georgian Ancient and Medieval Manuscripts is a National Centre for Manuscripts project that is underway, and the National Scientific Library started work on digitisation of its collections in 2008. A common element among all these programs is the keying-in of texts (often through volunteer work) rather than digitising to page images or using OCR.

Ghana. In 2006 the University of Cape Coast Library began digitisation in close cooperation with a number of external partners that either assist the library with training or support them financially. Available digitisation funds roughly equal the library's annual budget. The two main collections being digitised are Rare Books (for reasons of preservation and public access) and dissertations and theses (to benefit teaching, learning and research).

Kyrgyzstan. The Central Scientific and Medical Library of the Kyrgyz Republic has a $800 (USD) budget within the regular library budget ($17.000 USD annually). A recent project is to digitise all the materials related to the Kyrgyz surgeon Isa Konoevich Akhunbaev (1908 – 1975). The library also participates in the Republic's Development Centre for Public Health, which is an initiative to develop new, national health care related digital library information services. The Library already had experience, a scanner

and software in digitisation. The Akhunbaev project ran from June 2008 to December 2008 and scanned 2.800 pages from books, documents and photographic negatives. The primary objectives of the project were preservation and facilitation of research and public access by both the general public and subject specialists. Future projects will add other medical profession related heritage material and to create a collection on "Prominent Doctors of Kyrgyzstan."

Malawi. A digitisation project at Bunda College Library (BCL) at the University of Malawi included about 1.000 printed books, serials, documents and printed illustrations, with an average of 100 pages from each collection. Started in November 2007, the project ended in December 2008. The primary audience for the project were students, faculty, researchers and the general public, and the purpose was to preserve and extend access. The project support of $20.000 USD came from external funding, which was used partly for scanning equipment but largely to establish a centre of expertise at BCL, including training and implementation of Greenstone digital library software. Training included librarians and archivists both from Malawi and from Tanzania and Mozambique. BCL is now sharing its experiences with other libraries in Malawi and giving them support. New digitisation projects will be financed from local resources.

Poland. There are now nine regional digital libraries in operation throughout Poland. The first launched in 2001 with the Poznan Foundation of Scientific Libraries consortium, and became operational in 2002. The regional Polish Wielkolpolska Digital Library is the virtual door to content. The project includes 4.5 FTE staff and an annual budget of €200.000 for digitisation, computer hardware and software (which was written in-house in partnership with the Poznan Supercomputing and Networking Centre), maintenance, training and access. Participating libraries undertake the digitisation, with some using some high-end scanning equipment they already owned, and others using equipment from the consortium (including simple scanners for low cost digitisation). The purposes of the projects include preservation, wider access to cultural heritage, and creating resources for teaching and learning. Source material includes many types of printed or handwritten material. Continuous user feedback helps adjust content to meet their expectations. Thus far the Poznan consortium has scanned 10.000.000 pages over the past five years, and statistics on access show the collections to be in high demand.

Serbia. The Digital Library Department of the National Library of Serbia started its digitisation program in 2003. Financing to generate the working expertise came from a number of sources, including donations. As of 2008 there were 70 collections and 600.000 documents digitised. The primary purposes were preservation and public access to collections that were selected for their historic and cultural value. Source material included printed and rare books, handwritten documents, and sound recordings. Digital content is available on the web and in other formats such as CD's, with access open to both the general public and to Slavistics specialists and cartographers. The digitised material is in high and growing demand. The current library budget includes €25.000 for equipment and 7 staff who were redeployed from other library departments, and who work with volunteer scanner operators.

South Africa. The South African Music Archive Project (SAMAP) is being produced under the auspices of Digital Innovation South Africa (DISA), with some work performed in partnership with the International Library of African Music (ILAM) at Rhodes University. Digitised materials include sound recordings from analogue tapes and gramophone records. SAMAP provides access via the web to much of South Africa's music heritage to "promote multidisciplinary research in the field of popular music and culture" and to give open access to these hidden "politically sensitive or subversive" music treasures from the past new global audiences.

Tajikistan. The Rare Book Digital Collections by the Central Scientific Library of the Academy of Sciences of the Republic of Tajikistan is an ongoing digitisation program that serves two goals: preservation of the rare books collection and electronic access for research and teaching. Access is provided over the web using Greenstone software and on CDs. One particular collection digitised in 2004-2005 scanned 15.000 pages of rare book pages, photographs, maps, etc. The project included material from four different rare book collections, including a "Language and People" collection that digitised the 1899 published "Russian-Tajik Dictionary (which was the first linguistic Russian-Tajik Dictionary in Tajik linguistic history) and the 1902 published "A travers le Turkestan Russe" (which has a pictures representing the everyday life and history of the people living in this region). The total budget of $3.000 USD was externally funded by the U.S. Ambassador's Fund for Cultural Preservation.

Uzbekistan. In 2005 the Uzbekistan National Library embarked upon an in-house digitisation program for its entire collection of 20.000 rare books, manuscripts, and photographic prints, as well as its repository of 7.000 dissertations. At that beginning digitising was a new experience and equipment had to be acquired. UNESCO donated a high end scanner suitable for rare books and manuscripts. Today, the digitisation program has a 25.000 US$ budget and a separate line in the Uzbekistan National Library's annual budget of 800.000 US$. As of 2008 there were about 22.000 pages and 1.500 photos digitised from 1.100 rare books and about 500 dissertations. In the next phase of the project 30.000 more pages will be scanned. The purpose of the digitisation program was primarily preservation and expanded access by both subject specialists and the general public. Searching the online catalogue is open to anyone. Access to the digital content is free and unrestricted on the premises of the National Library's but regulated and fee-based from outside the library. Planning is underway to enable open access in the future to some collections. As a spin-off of the original project, the "Republican Center of Digitalisation" will be established to cover both the national library's digitising needs and those of other Uzbekistan libraries.

Digitisation Projects and Costs

Overview of Low Cost Digitisation Projects. The thirteen projects were diverse in cost factors, budget, staffing, expertise, equipment, staffing, and the amount of content digitised. In the table below, salaries "on project" means temporary staff for duration of project and on project budget; "re-deployed" means staff working at the library who are full- or part-time assigned to the project and for the duration of it.

Table 1: Project Duration, Budget, Cost factors and Results

Country	Library Annual Budget	Project Budget	Training needed at start project	Equipment etc. needed in project	Salary on project or re-deployed	Number of pages
Armenia 2008- 2010	280.000 US$	47.000 GBP	YES; cost= 1700 GBP	YES; scanners	re-deployed	15.000
Bulgaria 2004 -2006		23000 US$	outsourced	outsourced	outsourced	106667
Ethiopia ongoing from 2008	$1.4 million USD	$10.000 USD initial	NO	Heavy duty scanner	re-deployed	
Georgia 1998 - 2008; ongoing	From $4.5 million USD (NPL) to $210.000 USD (NCM)	Relatively small	YES	YES; scanners	re-deployed + volunteer work	Ongoing. 2008: 1.500 books NPL, 2.000 pages NCM, 1.000 pages NCL
Ghana ongoing from 2006	120.000 GHC	124.556 GHC	YES (costs of 12.000 US$)	YES; scanners	re-deployed	In the end over 500.00 pages from over 2000 items
Kyrgyzstan June 08 – December 08	17.000 US$	800 US$	YES; received for free	Scanner available; YES: digital camera	No costs; assigned from other project	2800
Malawi 2007- 2008	300.000 US$ (university total)	20.000 US$ (also setting up centre of expertise)	YES	One scanner needed; one available	re-deployed; receiving small incentive	1000 x 100 pages average
Poland Ongoing from 2002	Digitisation support is small percentage of 200.000 total consortium budget	Libraries scan and do so on their own budget	YES; free; covered by consortium	Scanners available; all else needed		Total for 11 libraries in 5 years: 10 million
Serbia Ongoing from 2003	3.200.000 Euro (25.000 set apart for digitisation equipment)	Many sources; cannot be calculated	YES	YES	Permanently re-deployed Scanner operators no costs (see story)	Ongoing: in 2008 600.000 documents (not pages!)
South Africa 2005 -2007 (phase 1)	DISA is independent body	400.000 US$ (SA R 4 million)	YES	Some present; sound digitisation	Salary on project	30.000 sound files + metadata

Country	Library Annual Budget	Project Budget	Training needed at start project	Equipment etc. needed in project	Salary on project or re-deployed	Number of pages
Tajikistan 2004 -2005	30.000 US$	3.000 US$	NO	YES	unknown	15.000
Uzbekistan Ongoing from 2005	800.000 US$ (20.000 set apart for digitisation)	25.000 US$	YES	YES (donated)	re-deployed	Ongoing: in 2008 22.000 and 1500 photos

Equipment, Facilities, Software Costs. Described below were the resources necessary to digitise collections (equipment, computers, software and other facilities), which expands upon the information contained in Table 1.

Table 2: Resources Needed, Available and Costs

Country	Scanner	Computer + Storage	Digital library	OCR software	Facilities & all else
Armenia	Two high end scanners = 40.000 GBP total	Graphic workstation; 2TB storage	No costs; Greenstone = FOSS DL	No costs; Abbyy Fine Reader; available	UPS for reliable electricity
Bulgaria	outsourced	outsourced		NA	
Ethiopia	HP Scanjet Scanner,	Server (8000 US$)		Needed, type x	Storage media, 200 US$
Georgia	NA; keying in mostly	NA	No costs; Greenstone = FOSS DL	not used	
Ghana	HP Scanjets 5590, 8270; ATIZ Book Drive			Adobe Professional 8, Omnipage	Extra furniture and shelves
Kyrgyzstan	available: HP 4600 ScanJet; needed: Digital Camera xx		No costs; DSpace = FOSS Repository	No costs; Abbyy Fine Reader; available	Furniture + shelves (no extra cost)
Malawi	Scanjet 5550C available; 5590 in project = 4700 SA Rand		No costs; Greenstone = FOSS DL	Lacking (an error in hindsight)	
Poland	High end available; computer workstations + low end scanners in project	Consortium for 11 libraries: in 2007 new server + discs = 36.000 Euro ; new upload software =10.000	In-house software writing	DjVU	
Serbia	High end Book Eye scanner = 33.000 Euro Epson GT 15000 = 1500 Euro	Storage + servers = 10.000 Euro BookRestore software= 1500 Euro	Commercial solution; planning for OAI access	Abbyy Finereader = 1000 Euro	
South Africa	Sound playing + digitisation available	PC's metadata assignment available	No costs; Greenstone = FOSS DL	NA	NA
Tajikistan	Needed, type x		CD and website access		
Uzbekistan	High end Plan Scan A3 Elar; in project but donated		IRBSIS (Russian); KARMAT (Uzbkstn)	Specialised OCR from "Computer Asia"	

Human Resources. Regardless whether a project uses specialised scanning equipment or a simple scanner, or is performed in dedicated office space with reliable electricity or in a spare room with power generator, every scanned page and bit of metadata must be created and added with human effort. These projects require team work to develop skills and expertise, which often is achieved through "re-deployment" of existing staff rather than attracting staff already skilled in digitisation work. This choice saves on separate salary costs but introduces training as cost factor. Tables 3 to 6 report on teams, training needs and costs related to training.

Table 3: Type and Number of Project Staff, and Time Devoted

In each cell, the number before the semi-colon is the number of people assigned to the project, and the number after the semi-colon is the percentage of time devoted to the project (u=unknown)

Country	Director	Metadata specialist	Curator	Digitizer (scanner operator)	Photographer	Technical support staff	Technical development staff
Armenia	1;5	1;7		2;30			1;2
Bulgaria		outscrd		outscrd		outscrd	
Ethiopia	1	1	1	3		1	
Georgia							
Ghana	1	1		2		1	1
Kyrgyzstan	1;5			2;70		1;15	1;10
Malawi	1;2	1;8		2;20			
Poland	1;100	1;100		5;u		1;100	
Serbia		1;50		2;100		1;100	1;100
South Africa	1;10	2;100		1;100		1;100	
Tajikistan							
Uzbekistan	1;20	1;50	1;50	2;50	1;50	1;50	

Notes:

- Georgia: volunteers key-in from original materials. No staff were dedicated or re-deployed
- Ghana: one Education Specialist and two Evaluation Specialists
- Malawi: Director and Metadata Specialist are the same person (the College Librarian)
- Poland: 100 % time on the project from consortium staff. The 5 Digitizers = total average number at work in consortium libraries. Technical development staff from the Super Computing Centre.
- Serbia: Digital Library Department + 3 "Indexing Operators" work for 100 % of their time
- South Africa's digitisation project requires more metadata specialists and fewer digitisers, which is the reverse of the text scanning situation.

Table 4: Training Needs

An "x" = training needed. "ntn" no training needed. "N/A" = not applicable

Country	Project manag e-- ment	Application of technical standards	Preparation & handling materials	Technical equipment operation	Post-digitization processes, e.g. editing	Metadata creation	Digital presser-vation
Armenia			x	x		x	x
Bulgaria	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ethiopia	x	x	x	x	x	x	x
Georgia	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ghana	x	x	x	x	x	x	
Kyrgyzstan	x	x	x			x	
Malawi		x	x	x	x	x	
Poland		x		x	x	x	x
Serbia	x	x	x	x	x	x	x
South Africa	ntn	x	x	x	x	x	x
Tajikistan							
Uzbekistan	x	x	x	x	x	x	x

Notes:

- Bulgaria: project outsourced
- Georgia: volunteers key in from original; no separate training needs
- Poland: needs identified participating libraries; training by consortium staff

Table 5: How Training was Organised

An "x" = training provided. "N/A" = not applicable. Note that "learning on the job" is a soft cost, as is "independent study." Using external consultants to provide in-house training or sending staff to external courses likely is a direct cost for the project.

Country	In-house, using project staff	In-house, using organization's own consultants	In-house, using external consultants	Attendance at external courses	Independent study	Learning on the job	Other
Armenia			x	x			
Bulgaria	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ethiopia	x	x				x	
Georgia	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Ghana			x	x		x	
Kyrgyzsta n							
Malawi	x		x			x	
Poland		x		x	x	x	
Serbia	x	x	x	x	x	x	
South Africa	x						
Tajikistan	Information to follow						
Uzbekistan	x		x			x	

Notes:

- Kyrgyzstan: external courses from the 'Kyrgyzstan Library Information Consortium"
- Bulgaria: project outsourced
- Georgia: work done by volunteers keying in from original: no separate training
- Poland: "in-house" means "by consortium staff," which must keep up with developments through "external course attendance" and "independent study"
- Serbia: a department within the national library, with advice, expertise and training more easily available than when work is done in single and smaller libraries
- Uzbekistan: "when members were receiving training there weren't good trainers, appropriate training programs and courses on digitisation in the republics"

Table 6: Staff Training Costs

Country	Type of staff trained; how many of them; information on costs
Armenia	Metadata specialist, digitizer. Soft cost: n the job. Direct: 1.700 GBP in-house by external consultant
Bulgaria	Not applicable; outsourced
Ethiopia	2 data encoders; soft costs only: trained locally
Georgia	Not applicable; keying in from original
Ghana	Operator, digitiser Meta Data Specialist, Technical Support: 9 people; soft cost: learning on the job at GHC 4000; direct costs: US$ 12.000 for external training
Kyrgyzstan	The operator, the manager of a network, the expert in marketing, the librarian: 4 people; soft cost 1: learning from operational experience in AUCA libraries, experience; soft cost 2: free of charge seminar in American University of Central Asia
Malawi	no separate cost information available
Poland	For the 11 libraries together per year: Digital Library software operators=20 people, Scanner operators=20, Cataloguing=10, Quality Assurance=20; costs: free of charge and training given by consortium . For the consortium: direct costs hiring external trainers: 1000 euro for DL software training, 1000 for Scanner training
Serbia	The project reports: "..In the Educational center of the NLS we had many courses and lectures of foreign experts on the topics of digitization (fees were covered by Library or by foreign donators), but the operational staff is trained in-house, face-to-face trainings with our specialists. The Department for the Development of the Digital Library has seven employees."
South Africa	Capacity building and skills development for all project staff and on project budget
Tajikistan	no separate cost information available
Uzbekistan	Staff involved in technical digitisation work; number not given; soft costs: learning on the job

Financing Digitisation Projects

Finances must be found before the digitisation project can begin. Many cost factors can be met through alternative means, such as staff re-deployment, learning from colleagues or on the job, and use of free and open source software. However, other costs (such as scanning equipment) involve out-of-pocket expenses (as described in tables 1 and 2 above). This equipment is a direct expenses that, when ordered from abroad costs, may need to be paid in foreign currency. After the initial purchase, there may also be costs for specialised maintenance (for high-end scanners) or replacement (for low end scanners). The choice of scanner or other equipment, such as for sound recording, is determined by the collection being digitised. Purchase of a less expensive scanner can result in the loss of quality or functionality, or could require more time. In some cases, less-expensive equipment may not be an option, such as for the digitization of rare books that are in poor condition or for projects with many pages or prints that must be scanned in a short time period of time. Almost all projects considered outsourcing the scanning process, but many reported that this option was too expensive or unavailable in the country, or that the collection was too valuable to allow it to leave the library or archive. Two projects that did outsource were in Bulgaria (which was then able to shift costs from equipment and training to financing the outsourcing) and Georgia (where most digitisation is done by keying in from originals rather than scanning and the work is performed by volunteers).

Table 7: Funding Sources at Project Start and Over Time

Country	Funding Source(s)	Funding Sustainability
Armenia	Grant from BL's "Endangered Archives Programme" (mainly equipment)	With scanner acquired, training received: continue on library budget
Bulgaria	UNESCO funding	
Ethiopia	Equipment and software from Association of African Universities	Indirect funding from regular university budget
Georgia	Small funding from different organisations: UNESCO, Open Society Foundation, US Embassy, other fundraising activities, e.g. local business support. Indirect financing through volunteer work.	
Ghana	The Worldbank TALIF Fund; Ghana Government Support	
Kyrgyzstan	Indirect funding by positioning the project in another program financed by the World Bank	No further information available
Malawi	Indirect funding by linking to a project building up a support network and starting centres of expertise (Greenstone DL's and digitisation); Koha Foundation and INASP funded	Scanners acquired, training received: "the project will continue with local resources"
Poland		(continuous financing stream for regional consortium of libraries)
Serbia	Library budget + fixed annual budget from the Ministry of culture for equipment + funding from OSI	Financed from the Library budget and special department and special staff
South Africa	South Africa National Research Foundation	Proposal for new funding completed
Tajikistan	One source is the US Ambassador Fund for Cultural Preservation	No further information available
Uzbekistan	One source is library's annual budget; another a UNESCO donation (equipment)	Will establish the Republican Center of Digitalisation to make digital collections for other Uzbekistan libraries

Results of Digitisation Projects and Access

The summary of examples from the eIFL.net countries explained the "why," "what," and "for whom" of the projects. Most of the projects reported that the most important reasons for digitisation were preservation, followed closely by easier and wider access. The target audience included the general public (to provide access to the historical and cultural collections), and students and faculty at schools and universities (to provide teaching and learning resources). While "number of pages scanned" reported in Table 1 is one indicator of the scale of accomplishment of these projects, more important are the measures below, including how access is provided, the ability to browse on a website, links from other resources (such as the online catalogue, or OPAC), and alternative forms of access (such as distribution on CD's).

Table 8 – Means of User Access to the Digitised Content

Country	Open access	Restricted access	Users have to pay
Armenia	thru OPAC		NO
Bulgaria		in-house users	NO
Ethiopia		for now in-house	NO
Georgia	Books	manuscripts	NO
Ghana	Partly	partly in-house	
Kyrgyzstan	thru OPAC & digital library		NO
Malawi	thru OPAC & digital library		NO
Poland	thru OPAC & digital library	some of the textbooks	NO
Serbia	thru OPAC & digital library	partly in-house	NO
South Africa	thru OPAC & digital library		NO
Tajikistan	on CDROM and web		NO
Uzbekistan	OPAC for discovery, not access	in-house users	Yes (if from outside library)

Table 9 – Level of Use of Digitised Collections (Selected Data Available)

Country	Level of use of digitised collection
Kyrgyzstan	The finished part (still ongoing) drew much attention last September at a "hundred years ago" memorial day
Poland	For 11 libraries and after five years since start: "statistics show that we have had 7.5 million visitors. On average we have 130-150 concurrent users"
Serbia	"Digital Library is one of the most successful project in the Library during last years. The number of users is growing rapidly, now we have more virtual users than in site users – physical users – 1000 per day, virtual users of our web site with all services – 7000 per day, 1.700 per day for Digital library .."
South Africa	Undetermined as yet
Uzbekistan	".. not a high level .."

When digitisation is of text then there are advantages in trying to not only scan to an image of that text but to also create so searchable full text. This can be created through Optical Character Recognition (OCR) software, with varying degrees of accuracy, or by manually re-keying the data from the original manuscript. In contrast to textual information, scanned images can be and found only through their metadata.

Table 10: Use of OCR or Manual Re-keying to Create Full Text

Country	OCR Package applied	Accuracy of result	Information on OCR process + costs	Keying-in applied
Ethopia	Omnipage		fully available in questionnaire	No
Georgia	No: does not handle Georgian script			almost exclusively
Ghana	Adobe Professional 8; Omnipage	90 %	not available	Keying-in also applied; causes delays
Poland	DjVu	various		No

Digitisation Projects and Standards

Digitisation must certain standards to provide useful and true digital representations of the physical original. “Useful” means that the scanned images of print material can be used for other purposes, such as to convert the full text into searchable data through OCR. “True” representation means that the image should accurately represent the original, e.g., researchers studying a digitised copy of a very rare handwritten illustrated manuscript should be able to get an image that is equal to the original. There are a number of published guides and guidelines that to help ensure “true and useful digital representations.”

Table 11: Types of Guides or Guidelines to Digitise Documents

Country	Used guidelines?	If yes: which guidelines
Armenia	YES	IFLA Digitization Guidelines
Bulgaria	(outsourced)	
Ethiopia	NO	
Georgia	NO	
Ghana	NO	
Kyrgyzstan	NO	
Malawi	NO	
Poland	Own research	Published own guidelines as consortium (in Polish)
Serbia	YES	IFLA and UNESCO Guidelines, PULMAN guidelines
South Africa	Own research	DISA guidelines based on international standards and best practice and adapted for use in sound digitisation
Uzbekistan	NO	

Sustainability of Digitised Collections and Access

The sustainability of digitised collection and access requires careful attention and a sufficient budget and ongoing work to ensure continued access. Long term access requires a digital preservation strategy for timely migration to new media, upgrading of metadata as standards may change, and continuous review to ensure that the needs of different user groups are being met as technology and demands change.

Table 12: Provision for Future Updates

Country	New materials digitized	New metadata added	User Interface changed	File Formats changed	Digital Preservation Strategy
Armenia			x	x	Yes
Bulgaria	permanently	permanently	x		No
Ethiopia	x		x	x	Yes
Georgia					No
Ghana	(not answered)	(not answered)	(not answered)	(not answered)	(not answered)
Kyrgyzstan	x	X	x	x	Yes
Malawi	as they arrive	X	x		No
Poland	permanently	permanently	with new versions	x	Yes
Serbia	monthly	Monthly	every year	occasionally	N ; will be developed
South Africa	x	X	x		Yes
Tajikistan	(not answered)	(not answered)	(not answered)	(not answered)	(not answered)
Uzbekistan		X		x	No (under development)

Conclusion

Many libraries in eIFL countries have undertaken or are planning digitisation projects. This report on case studies of low cost digitisation projects brought together some of their experiences. Underlying this report are the full answers to each of the surveys, each forming a mini-report by itself. This Final Report is simply a selection, annotation and interpretation of this original material. Based upon the data provided by the reporting institutions, the following tentative conclusions can be drawn:

- For the most part, in-house digitisation projects are relatively recent undertakings, but they seem successful and here to stay. Digital libraries increasingly will be created with digital content coming from the collections owned by the libraries.
- There seem to be two motivations for digitisation: preservation of physical collections in poor condition, and the wish to provide users with more readily accessible local digital content. The value of the latter can proven by usage statistics. Some of the cases also demonstrate that these two functions can both be achieved effectively through the same project.
- A digitisation project often does not have to "break the bank," but there are some types of material do require equipment that is so expensive that only outside funding or donation can kick start the project.
- Kick-starting a project requires the library to have sufficient funds and human resources, or to find creative ways to cover costs. For the provision of some specific equipment, it may be valuable to investigate the purchase and installation of centralised, streamlined equipment or to find other ways to make this funding and equipment available.
- An attractive and resourceful way to get more digitisation done or to help a project have an easier start is to foster collaboration among libraries through consortia and by creating centres of expertise to provide training and support or lend equipment to first-time digitisation projects.

Appendix A
List of eIFL Countries Contacted

eIFL country	Included in case study (or reason for not being included)	Participant
Armenia	Yes	Fundamental Scientific Library
Bulgaria	Yes	New Bulgarian University + others
Ethiopia	Yes	Addis Ababa University Comp Centre
Georgia	Yes	Nat Parliament, Manuscripts, Scntf Lib
Ghana	Yes	Univ of Cape Coast Library
Kyrgyzstan	Yes	Central Scntf and Medical Library
Malawi	Yes	Bunda College Library, Unv of Malawi
Poland	Yes	Wielkolpolska Digital Library
Serbia	Yes	National Library of Serbia
South Africa	Yes	DISA
Tajikistan	Yes	Central Scientific Library
Uzbekistan	Yes	Uzbekistan National Library
Mali	No. Born digital materials only.	Faculte de Medecine, Pharmacologie
Azerbaijan	No	
Belarus	No	
Kenya	No (still in planning stages)	
Macedonia	No	
Palestine	No (still in planning stages)	
Senegal	No	
Sudan	No	
Lithuania	No (projects above “Low Cost” criteria)	
Moldova	No (projects above “Low Cost” criteria)	
Slovenia	No (projects above “Low Cost” criteria)	

Notes:

- “No” means “No Response”.
- Countries marked “No (still in planning stages)” are countries not yet involved with digitisation (to the best of our knowledge) but are planning to do so.
- Countries with “projects above ‘Low Cost’ criteria” have digitisation projects on relatively high budgets or funding schemes and were not included in the study, but they voluntarily responded to the initial Call.

Appendix B
Frequently Asked Questions (FAQ) About Low Cost Digitisation Projects

Q: I have a library collection that is physical but want it digital: where do I start

A: The first step is to educate yourself about some of the issues. You may find the following digitisation guidelines to be helpful:

- IFLA & UNESCO http://www.ifla.org/VII/s19/pubs/digit-guide.pdf
- PULMAN http://pulmanweb.org/DGMs/DGMs.htm
- DISA http://www.disa.ukzn.ac.za/index.php?option=com_docman&task=cat_view&gid=62&Itemid=88
- the NINCH Good Practice Guide [NINCH] http://www.nyu.edu/its/humanities/ninchguide/index.html
- the iMARK Information Management Resource Kit Module "Digtization and Digital Libraries" <http://www.imarkgroup.org/moduledescription_en.asp?id=1> Also on CDROM in a number of languages]
- "From paper to collection" (Loots et al., 2004). http://greenstonesupport.iimk.ac.in/Documents/FromPapertoCollection.pdf

Talking to other librarians who are geographically close to you and share the same language should also be very helpful.

The next step will be to develop a plan of action. For example, the iMARK Module helps provides many useful exercises and checklists.

Finally, get started. The most important lesson learned from all of the projects we studies is this: if you can finance the cost of your scanner and manage the initial training, apply that knowledge as soon as possible so you can learn by doing.

Q: How can I finance my digitisation project?

A: Most of the projects we studied where financed through one of these sources: (1) obtaining external funding (e.g., grants, foundations, special ministry funding) for scanning equipment and IT-related investments, (2) handling staff costs through redeployment or volunteer work, or (3) additional funds from the regular budget to support training, (4) asking colleagues at other nearby institutions to provide training.

Q: Do most libraries work together on digitisation or by themselves?

A: The experience varies by country. For example, Poland demonstrated the advantages of regional cooperation, which enabled costs to be shared, central coordination of arrangements, and centralize access to the digital library. Uzbekistan is seeking to establish a digitisation centre at the national level, which would be of a benefit to all libraries. In South Africa, DISA is an example of a group that was able to share expertise, project management and hosting. Nonetheless, in many of the other projects the libraries worked on their own. Some additional information about the participation in international programmes is available from the British Library's "Endangered Archives."

Q: I know which of my collections I should digitise. Should I outsource this work or do it in-house?

A: Most of the projects we studied chose to do the work in-house. The most frequently cited reasons for doing so were because: (1) there were no outsourcing facilities in their country; (2) the collections being digitised were too fragile or special to allow them to leave the library; (3) the out-of-pocket costs for outsourcing were greater than what was available in the budget; (4) although the work might take longer, it was possible to extend the project over time and use in-house staff to complete the project. There were two instances of outsourcing in the study. Bulgaria outsourced the digitisation of the journal collection scanning, and Georgia outsourced to volunteers the work to type and input the text of old manuscripts.

Q: What is OCR? What does it do and should I try it? Are there alternative solutions?

A: After scanning a page, Optical Character Recognition (OCR) software can be used to read the scanned text images and produce full text as if it were typed. The resulting full text enables indexing and searching through the text. However, OCR usually can read only fairly standardized type fonts, so it cannot always handle local fonts or handwritten material. Even when the type font is standardized, the accuracy of the conversion will vary depending upon the quality of the OCR software. The use of OCR also requires additional steps to convert the scanned image to text and to proofread it, but without OCR the only text that will be searchable will be the metadata that the library might generate. Examples in the study where the library used OCR software were Ethiopia, Ghana and Poland.

If you wish to generate full-text, the alternative is to re-type the text (or to outsource that work to another organization). Georgia reported their experience with keying in of rare books by knowledgeable volunteers. This process can be very time-consuming and can also introduce the opportunity for error. You can sometimes outsource this work (particularly for English language text material), but no outsourcing of this was reported in the survey.

Q: Which equipment and software do I need?

A: The table on the next page provides an overview from the countries that participated in this study in terms of the materials they digitised, the approach they took and the equipment and IT infrastructure they employed.

Country	LCD case study participant	Material type	Approach	Equipment & related IT
Armenia	Fundamental Scientific Library	Early printed books and periodicals	In-house using a high end scanner. Participated in the British Library "Endangered Archives" program	High end ICAM GUARDIAN Book Handling System. Phase One P45 Digital Camera Greenstone FOSS for access
Bulgaria	New Bulgarian University + other universities, cities, & public libraries	Serials collections	Outsourced to a Bulgarian company	
Ethiopia	Addis Ababa University Computer Centre	All Ethiopian collection materials centrally held by the university	In-house. Used OCR after image scanning	HP Scanjet Large server
Georgia	National Parliament Lib, Nat Manuscripts Lib, Nat Scientific Lib	Ancient and Medieval Manuscripts, Modern Georgian Writers, Scientific Library print collections	Re-typed text on a large scale using volunteers	Greenstone FOSS for access
Ghana	University of Cape Coast Library	Rare Books, Dissertations and Theses	In-house. Mixture of scanning images, OCR, and re-typing text	HP Scanjets 5590, 8270;ATIZ Book Drive OCR: Omnipage
Kyrgyzstan	Central Scientific and Medical Library	All materials related to the Kyrgyz surgeon Isa Konoevich Akhunbaev (1908 – 1975)	In-house	HP 4600 ScanJet Digital camera DSpace FOSS for access
Malawi	Bunda College Library at the University of Malawi	Malawiana: printed books, serials, divers documents, printed illustrations	In-house	HP Scanjets 5550C, 5590 Greenstone FOSS for access
Poland	Wielkolpolska Digital Library	Almost any type of printed or hand written material	Consortium approach	High + low end scanners DjVU for OCR
Serbia	National Library of Serbia	Printed (rare) books, handwritten documents, and some sound recordings	In-house. Used mid-range and high end scanners. Employed army conscript volunteers as scanner operators	High end Book Eye scanner; Epson GT scanner Abbyy Fine reader for OCR
South Africa	DISA	Sound recordings (open reel tapes, gramophone records)	Centrally managed projects, with central hosting of content	Sound digitisation hardware + software Greenstone FOSS for access
Tajikistan	Central Scientific Library	Rare books, photographic illustrations, maps and others	In-house. Undertook thematic scanning projects	
Uzbekistan	Uzbekistan National Library	Rare book collection (also photographic prints), all Uzbekistan dissertations	Developing into a centre of expertise for the country	High end Plan Scan A3 Elar