Bożena Bednarek-Michalska
Nicolaus Copernicus University Library in Torun
Poland

**Kujawsko-Pomorska Digital Library (KPBC – kpbc.umk.pl):
the pragmatics of creating a digital library**

Abstract

The article talks about developing a practical model for creating a digital library, which is still quite an experimental activity in Poland. Based on her team's experiences in creating the KPBC, the author discusses steps to project planning and completion, such as, defining project components, objectives, foundation, team management issues and criteria for team selection. Furthermore, the author deals with planning and supervision of work processes, provides an outline of recommended procedures for digitalization and shows an example of a project budget. Thanks to the exchange of practical knowledge on creating digital library projects, future undertakers of similar ventures can find their mission more manageable.

* * *

The pragmatics of creating digital libraries has not yet been thoroughly described around the world, because everything associated with the development of digitalization is changing dynamically and our understanding of this phenomenon is still insufficient. We have some experience in developing electronic collections in Poland, maybe not as extensive as in the United States, Australia or other European countries, but overall, digitalization efforts remain quite experimental. Due to this, librarians around the world make efforts to describe the pragmatics of creating digital resources based on their own experience, to support one another and to validate the chosen methods. As a common practice, the creation of models produces standards to be met in order to obtain a high-quality final product. *Collection strategies and practices are not yet fully developed to take account of these changing circumstances, nor are their legal, organizational, and business implications fully understood.* [1]. Many such strategies and good practices can be found on websites of library federations, consortia or associations and they should be consulted when deciding to embark on a digital library project.

In spite of having such slight experiences in this area, we should share our practical knowledge to make the next Polish digital libraries better than the previous ones. Hence, the intention of this essay is to share the experiences of the Torun and Bydgoszcz librarians, which enabled the creation of a regional digital library. This article is the second of two, which discusses the standards [2] and pragmatics of the Kujawsko-Pomorska Digital Library (KPBC).
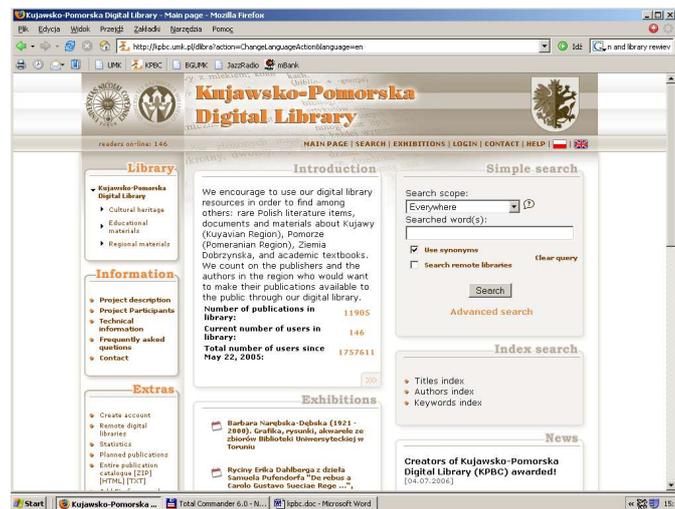
The Kujawsko-Pomorska Digital Library (KPBC).
Kpbc.umk.pl

Project planning

To begin, one essential point is that, in order to develop a good digital library, it is necessary to consider all elements and discuss them in a larger group. It is vital to master the knowledge on digitalization and take into consideration the strategic planning of the institution. In my opinion, the key quality in digitalization is the skill of planning and managing of, not only teams, but also, tasks and information and, finally, the skill of supervising the implementation of those tasks. Anybody wishing to start a digital library project has to bear in mind that it is a costly venture and it should be developed in such a way that will allow the results to be enjoyed over many years, which involves applying international standards. At the conference in Bielefeld, Germany [3] [4], it was stated that the life expectancy of the currently created digital resources amounts to 20 years. Some specialists claim that it is even shorter. What will happen after? These are the problems European countries have to solve and we should be aware of them to learn how to manage them in the future. Therefore, it is so important to meet the international standards and guidelines developed by specialists.

Project planning is an extremely important skill we have to acquire in order to become professionals and to get rid of the false thinking that 'we will manage somehow.' Before beginning, it has to be acknowledged that an unplanned and chaotic approach, sooner or later, will produce an incoherent and unfeasible project. For example, the project may fail due to some undefined initial requirements that have to be met. In developing the Kujawsko-Pomorska Digital Library, we made great use of the knowledge and skills of our team. This critical knowledge can be acquired in many ways, for instance, in the course of university management studies, where students realize the importance of planning and designing, why everything has to be recorded, what is a well-structured plan, what is a business plan, a work plan and efficiency indicators, or how to obtain the best final results.

Moreover, it is important to apply previous practical experience gained in projects within non-governmental organizations. After completing five projects, each following one seems easier, even if it differs in terms of objectives, scope, importance and complexity. Another opportunity to learn is a first-time preparation of an application for EU funding.

Unfortunately, not all project components can be developed in a professional way. Before embarking on such an endeavor, it is advisable to precisely define the audience of the digital library and analyze its needs and requirements, an almost impossible task to achieve in Poland as decision-makers react with surprise while lack of time usually paralyzes the process. If there were current assessments of digital library users and digitalization research

and if there were statistics on the Internet or new technologies, it would be much easier to prepare such projects. However, regrettably, this is not the case in Poland, where there are no research studies necessary to modernize our library services. These usually have to be ordered especially for the purpose of a given project, but it takes such a long time to obtain results that waiting is not always an option. Such research is made on a regular basis, for example, in Germany (BIBWEB), where before a project is launched, potential users are questioned if they really need such services and then, based on the analysis, a draft of the project is prepared.

Project components

Below are some key components of the project, usually sketched out before starting the tasks. However, it should be kept in mind that not all of the mentioned elements are necessary. It all depends on who finances the project and what requirements the financing body sets forth. It also may depend on the project level, whether local or international, and on which partners are involved in the project's completion.

Introductory information:
Information about the institution;
The potential of the university and the region;
The potential of research libraries of the region.
1. The goals and the premise of the Digital Library.
The contribution of the project to strategic planning of the country and the region.
2. Project description.
The origins of the Kujawsko-Pomorska Digital Library (KPBC);
The institutions involved in the project;
Project location;
The premise of the project;
The setting of KPBC and its relation to other Polish and regional projects;
Project promotion;
Technologies.
3. Project justification.
Risks to the realization of the project.
4. Detailed goals and tasks of the project for 2003-2006.
The tasks and work plan;
Long-term benefits resulting from the project's implementation;
Expected outcome – indicators.
5. Project management.
The management team;
Task management;
Evaluation and control.
6. Financing the project.
Costs *(in thousands PLN).*
7. Post-completion strategy (operation of the project after its completion).
8. The future of the project for 2006-2008.
9. Bibliography.

As shown in the above outline, planning should begin from defining objectives and preparing a general description of the project. In the case of KPBC, these were described in the following way:

General goal: Building an infrastructure for the regional electronic platform, known as the digital library, that will provide quick access to knowledge (books, articles) while preserving the valuable documents of the region and the literary heritage (manuscripts, old prints), thus supporting the development of the intellectual and innovative potential of Poland.

Specific objectives:
1. *Creating 3 digitalization centers in the region;*
2. *Digitalization of written documents (approximately 20 thousand documents in the first phase, i.e. by the year 2008);*
3. *Opening access to resources on the Internet (new central electronic platform).*

The project description should include all of its key components and its premise.

The project premise consists of building a digital library composed of the following three collections:

1. **Scientific-Educational**, consisting of digital copies of selected academic textbooks, monographs and scientific articles.
2. **Cultural Heritage**, containing digital copies of the most valued items, such as, incunabula, old prints, manuscripts, and iconography collections.
3. **Regional**, consisting of digital copies of leaflets, posters, invitations, exhibition and fair catalogs, as well as, other materials from the region.

While defining the foundation of the digital library, we can specify the combination and selection of materials, digitalization and metadata standards, as well as, address copyright terms. We can also determine the level of access to resources (open or restricted) and identify its users, the scope of the project (regional or national), as well as, what technologies and software will be used. If other institutions are participating in the project, in addition to describing our institution's potential, that of the partners should also be addressed.

Following the description of the premise, we proceed to providing the rationale behind building the digital library by giving a clear justification, at the same time, acknowledging the risks involved in the completion of our project. Next, we can outline the specific tasks and deadlines for completing them by preparing a work plan, such as below.

		Approximate deadline
1.	Beginning the project	January 2005
2.	Upgrading the digitalization centers	March 2005
3.	Purchasing equipment and furniture	April/May 2005
4.	Team training	Spring/Summer 2005
5.	Software implementation and testing the platform	June/July 2005
6.	Digitalization and data processing	September 2005/November 2006
7.	Launching the project online	September 2005
8.	Planned deadline for completing the project	December 30, 2006
9.	Planned deadline for obtaining the expected project results and settling the financial matters	November 20, 2008

Table 1: Project schedule. Milestones – stage 1.

Project and team management

In order to build a digital library, it is necessary to hold an idea, even a vaguely defined one, as well as, to have a proper team. It is particularly essential to have a leader who will persuade the team members to do the project and who will also lead them through it. The person in this position has to prepare a detailed action plan and select a team of competent members since undertaking such a project involves too much work for a single person to handle. Additionally, the leader should have exceptional managing skills since the coordination of tasks is highly complex and requires imposing rigorous discipline that will drive the project to its completion.

One extremely important element supporting a sound project can be international cooperation and implementation of international projects. By working together with more experienced colleagues, we acquire knowledge (know-how, know-what) necessary to accomplish our own tasks, at the same time, building up courage and getting rid of low self-esteem, which is an important motivator. The experience also helps in managing the team. The KPBC team is a large one, consisting of members from three regional libraries, where each member has been assigned precisely defined tasks. Although strictly imposed, the tasks can be modified. Since KPBC is a well-financed project and the team is paid according to clearly established rules, the work proceeds relatively quickly.

Criteria for team selection:
- Experience in library work;
- Knowledge and skills;
- Enthusiasm for work;
- Attitude towards innovations and technology;
- Confidence;
- Communication and teamworking skills.

The team must have a well-designed work plan. It is necessary to examine and discuss work processes, to improve them and make every effort to define procedures.

The team of the Nicolaus Copernicus University (NCU) consists of:
1. Project coordinator – administration, financing, cooperation with participating institutions, promotion, professional and quality supervision, and copyright/legal negotiations.
2. Assistant coordinator – supervision of the technical side of the project, equipment, software, agreements, bidding, standards.
3. Administrator – project documentation, finances, reporting, personnel matters, correspondence and other tasks.
4. Editor – editing the digital library, resource management, coordination of data entry and processing teams.
5. IT specialist – software, supervision of dLibra, technology solutions, statistics.
6. Chief selection specialist – selection of documents for digitalization, selection of materials from special collections, coordination of tasks.
7. Chief cataloger – bibliographic description, metadata, improving the overall resource, standards.
8. Technicians – digitalization, lab supervision, scanning standards, archiving of items.

The above makes up the main core of the KPBC team. Apart from the mentioned, there are additional 10 members at the NCU involved in the selection of resources, preparation, scanning, image processing (OCR) and other technical and organizational tasks. For many of them, these are completely new challenges and tasks that may generate interest and eagerness, but also, stress. In addition to the librarians, works on dLibra, servers and network involve specialists from the Poznan Supercomputing and Networking Center and the Nicolaus Copernicus University IT Center. The NCU also employs a special team of 3 persons responsible for managing the administrative side of structural funds, who assist the librarians in their administrative and financial tasks. Before the official opening of KPBC, it was calculated that, within the year 2005, the project involved 60 people from 3 Bydgoszcz universities and the provincial government. It was a really large team and a huge undertaking. The involvement of our partnering institutions is not described herein because in two Bydgoszcz libraries, the work is performed in smaller groups and the procedures are quite different, and therefore, should be addressed separately.

Work process planning

After selecting the team and determining the scope of responsibilities that can be assigned to each member, we can proceed to mapping out the work processes. We create a general

description of tasks to be carried out in order to fill the digital platform with resources and we assign particular persons to each task.

Work process elements:

- Selecting and qualifying documents for digitalization according to the initially adopted guidelines (selectors);
- Creating a table of items to be digitalized (selectors);
- Queuing the documents up for digitalization and controlling the queue (editor);
- Technical preparation of documents and their delivery to the laboratory (selectors);
- Scanning and archiving in the TIFF format (technicians);
- Processing of files created in the process of OCR and DjVu scanning (IT personnel, technicians);
- Creating a bibliographic record in the local Horizon catalog in the MARC21 format, converting the record into dLibra using DublinCore (catalogers);
- Publishing digitalized material (editor);
- Verifying the metadata in dLibra and Horizon, data correction and resource improvements (chief cataloger).

If we want team members to know exactly what and when tasks should be completed, it is worth including them in the preparation of the detailed work procedures, that is, the steps to be performed to obtain the expected final product.

Stages of document digitalization

The procedures recommended at the NCU Library in Torun

1. **Qualification of documents.**
   a) Identifying the documents, determining the methods and time of digitalization;
   b) Creating tables of up to 15 items with precise processing instructions (scanning quality, color and other);
   c) Saving the tables onto a shared disk available to all network users (using an agreed file format: graphics_003_doc);
   Performed by: librarians responsible for the respective collections.
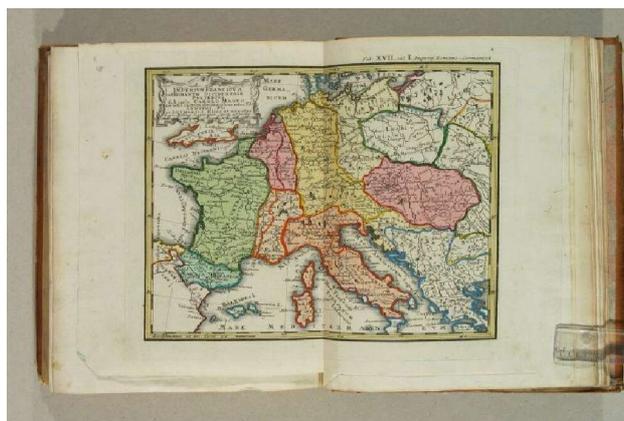
Old printing book in the digital laboratory, © Piotr Kurek

2. **Queuing of documents presented in tables.**
   a) Queuing the tables, prioritizing the sequence of documents to be scanned;

b) Informing the laboratory about the queue and its readiness for scanning;

c) Sending the tables to the digitalization laboratory;

d) Ongoing supervision of the digitalization process and observation of procedures.

Performed by: KPBC editor.

3. **Scanning and archiving.**

a) Ordering documents to be scanned by contacting the person signed under the table;

b) Signing the table for the librarian delivering the materials treating the table as an order slip;

c) Scanning documents in the lab according to tables set in a queue (remarks on scanning results should be sent to the editor);

d) Processing the results of scanning using the software provided with the scanner (image correction, framing etc.);

e) Saving files under a standard name (for many files, the call number corresponds to the file name or catalog), creating call numbers for boxes containing the medium with the digital version of the document; physical description of the digital version of documents (file format, medium type, date of recording, resolution, color details etc.);

f) Sending the prepared files for further processing;

g) Delivering the original documents to the technical department using the table as a receipt, as in section b).

Performed by: Digitalization Laboratory employees.

4. **Processing of files created in the scanning process.**

a) Obtaining file formats (djVu, html, pdf and others) indicated in the qualification and preparation process, OCR for some items;

b) Delivering files for further editing.

Performed by: Digitalization Laboratory employees, IT specialists.

5. **Publishing the files on the digital platform (dLibra).**

a) Linking the item to its record;

b) Loading files into one or several dLibra collections;

c) Supplementing the dLibra record by metadata from the digital version.

Performed by: KPBC editor.

6. **Creating a bibliographic record.**

a) Cataloging in Horizon;

b) Uploading bibliographic records into dLibra planned publications, into final collections;

c) Continuous quality improvements and verification of records in dLibra;

d) Making corrections in dLibra, browsing indexes for errors;

e) Returning documents to the storage facility or reading-rooms, as instructed in tables.

Performed by: librarians responsible for the specific special collections.

7. **Verifying metadata and uploading bibliographic records into the NuKat national catalog.**

a) Supplementing the existing NuKat records with new information on the digital version of documents;

b) Creating new records in NuKat with links to KPBC digital documents;

c) Linking existing records in NuKat to dLibra;

Performed by: chief cataloger.

The Digital Laboratory, © Piotr Kurek

It is important to remember that some work processes are performed concurrently, whereas others consecutively. It should also been foreseen that some processes will be held up and will not be resumed without additional efforts or costs. Once the tasks are assigned, it is vital to properly supervise their completion; therefore, it is also crucial to define the scope of responsibilities of every staff member.

Supervision of work processes:
- Supervision of work processes and finances, negotiations with authors and publishers – project coordinator (1);
    - Supervision of project documentation – project administrator (1).
- Supervision of hardware and software – assistant coordinator, IT specialist (1);
    - Supervision of the digital platform – IT specialist, i.e. dLibra administrator (1);
    - Supervision of scanning – chief technician (1).
- Supervision of uploading the resources and their quality – editors (3);
    - Supervision of resource selection – special collection expert (1);
    - Supervision of metadata – chief cataloger (1).

Sound planning of all essential steps to building the digital library and proper management of work processes will produce better results, thus convincing the employees of the significance of their work. It is necessary to continuously motivate the librarians by illustrating how they are pioneers of crucial changes for user services, and how they are indispensable in the advancement of their institution. By instilling the idea that they have started a lengthy scanning process, which will serve users for a very long time, the quality of their performance will improve. Furthermore, it should be emphasized that their efforts are contributing to new types of services connected to the digital library, such as, e-learning.

An example of a project budget

Category of expenditure	2004	2005	2006	Total
Qualifiable expenses				
Project administration	4,500	5,000	3,000	**12,500**
Project preparations (research, survey reports)	5,000	0	0	**5,000**
Preparation of documentation for the bidding	1,000	1,000	0	**2,000**

Renovation of rooms for digitalization laboratories	31,155	0	0	**31,155**
Installation of basic systems (computer network, power supply, air conditioning)	11,811	0	0	**11,811**
Anti-burglary system (windows, doors)	11,700	0	0	**11,700**
Equipment for digitalization laboratories	497,843	0	0	**497,843**
Special furniture for digitalization laboratories	9,594	0	0	**9,594**
Software and configuration	62,000	6,000	0	**68,000**
Digitalization	0	18,600	18,600	**37,200**
Project promotion	1,000	1,200	1,000	**3,200**
Staff training	3,000	3,000	0	**6,000**
Building a uniform web platform	3,000	2,000	2,000	**7,000**
Data processing	0	38,750	38,750	**77,500**
Unqualifiable expenses				
Value Added Tax (VAT)	164,617	1,540	0	**166,157**
Taxes on work assignment contracts	0	35,150	35,150	**70,300**
Total project cost	**806,220**	**112,240**	**98,500**	**1,016,960**

Table 2: Planned expenses (PLN)

The future of the project

Building a digital library is a costly endeavor and its completion should be planned to stretch over several years. For this reason, it is so vital to consider whether the project stands a chance within the institution and whether there will be adequate human and financial resources for its long-term maintenance. It cannot be stressed enough that the emergence of digital libraries and other electronic services requires us to modify the organization of library work and the very structure of our institutions. New tasks generate new work processes and, as a result, demand organizational changes. Tasks should be assigned and adjusted to the already performed ones or new teams should be formed to carry them out. For instance, creating new metadata for a digital library can be incorporated into the operations of the descriptive cataloging or subject cataloging departments, but who within our library ranks can be assigned the chore of editing the digital library? The library director is faced with a question of whether to create a new department to fulfill the needs of the digital library and perhaps other electronic services, or to form new positions in the existing departments?

Furthermore, new technologies being introduced in libraries produce identical problems. The NCU Library in Torun transformed its reprographic services laboratory with three full-time employees into a digitalization laboratory by simply buying the necessary equipment, training the staff and giving a new quality to work processes. However, not all libraries were fortunate to have such great employee potential. For example, our partners from Bydgoszcz were forced to start from scratch and find specialists to carry out the new tasks.

Present-day libraries, especially the larger ones, offer a growing number of electronic services, and thus need to campaign for more external funds for their development and to recruit competent IT specialists, photographers, economists, administrators to meet the new demands. Therefore, in addition to organizational changes, there are many issues related to recruiting new employees, motivating the old ones to adapt to the new chores, training etc. All these concerns should be approached with calm and careful planning in order to ensure the most advantageous solutions.

Conclusions

The above-described experience of the Torun librarians with building the Kujawsko-Pomorska Digital Library surely does not cover all the issues, which perhaps would have been of interest to other librarians, however it brings up some valuable points. This is not to

say that the model presented herein is the only way to create a digital library. Instead, there are as many possible ways as there are libraries, institutes or archives willing to follow. The more we share our methods of creating digital libraries, the better for the library community and the users. Perhaps this activity will produce a list of the most essential, as well as, secondary ingredients needed to develop good electronic resources in Poland. By sharing this experience, I hope to encourage our library community to exchange knowledge and make good use of the experiences of other digital collection teams. Let's form good practices, let's cooperate and the effects will come sooner or later.

Much can be learned from the solid approach towards work of our American counterparts, the simplicity and pragmatism, with which they move through activities. Instead of over-discussing how a digital library should be done, they submit themselves to the experience and gain knowledge through practice, they learn, record good examples and rules, create models and make them readily available on the Internet, and finally, they are open-minded and always willing to lend a hand. It is thanks to the experience of our American colleagues, the leaders of change, our path towards creating our own digital resources is shorter and easier.

Bibliography

1. The Digital Library Federation. Collect, Produce, Preserve, Use, Build. Accessed at: http://www.diglib.org/collections.htm (March 22, 2006).
2. BEDNAREK-Michalska, Bożena. Kujawsko-Pomorska Biblioteka Cyfrowa a standardy, *Biuletyn EBIB* [digital document], no. 4/2006 (74). Accessed at: http://www.ebib.info/2006/74/michalska.php (March 22, 2006).
3. 8th International Bielefeld Conference 2006. New Paradigms for the Digital Age [digital document] 07.02-09.02, Bielefeld University Library. Accessed at: http://conference.ub.uni-bielefeld.de/ (March 22, 2006).
4. BEDNAREK-Michalska, Bożena. Sprawozdanie z konferencji w Bielefeld, *Biuletyn EBIB* [digital document], no. 2/2006 (72). Accessed at: http://www.ebib.info/2006/72/bielefeld.php (March 22, 2006).
5. CZYŻAK, Dominika. Kujawsko-Pomorska Biblioteka Cyfrowa – stan zaawansowania realizacji projektu ZPORR, *Biuletyn EBIB* [digital document], no. 9/2005 (70). Accessed at: http://ebib.oss.wroc.pl/2005/70/czyzak.php (March 22, 2006).
6. A Framework of Guidance for Building Good Digital Collections [digital document] (2004). 2th Edition. Bethesda: NISO Framework Advisory Group. Accessed at: http://www.niso.org/framework/Framework2.html (March 22, 2006).
7. Guidelines for using resource identifiers in Dublin Core metadata and IEEE LOM, OCLC, [online]. Accessed at: http://www.ukoln.ac.uk/metadata/dcmi-ieee/identifiers/ (February 20, 2006).
8. WERLA, Marcin. Biblioteka cyfrowa jako repozytorium OAI-PMH [digital document]. Accessed at: http://dlibra.psnc.pl/index.php?option=com_content&task=view&id=62&Itemid=62&lang=pl (March 15, 2006).

Translation: Marta Sobieszek