# Czech Digital Library

*Bohdana Stoklasova; National Library of the Czech Republic; Prague; Czech Republic*

## Abstract

*The Czech Digital Library is a new national plan to integrate different library repositories in the Czech Republic with repositories of other cultural heritage institutions. The Czech Republic is a small country in the heart of Europe with a long tradition and remarkable results in the area of digital archiving that are recognized worldwide: in 2005 the National Library of the Czech Republic was awarded the first UNESCO/Jikji Memory of the World Prize for its contribution to the preservation and accessibility of its documentary heritage. Large national projects have been functioning for more than ten years, combined with other regional, subject- oriented, institutional, and other projects. These more or less isolated projects have produced millions of digitized pages accompanied by metadata. Now we face the challenge of closer cooperation and coordination of these projects under the framework of the Czech Digital Library. This paper deals with both the results achieved in the three large national projects and our concept for the future.*

## Introduction

The Czech Digital Library conceptualizes a new national integration of the different library repositories in this country with repositories of other cultural heritage institutions. The main goal of this concept (as stated by the Czech Ministry of Culture), which must be approved by the Czech Government during the upcoming months, was formulated as the urgent necessity of preserving and making permanent access available to collections of traditional, digitized, and digitally created documents as an important part of our national cultural heritage. The planning especially stresses the need for changed legislation, reorganization, a new assignment of duties among different institutions, and building the necessary infrastructure.

The Czech Republic is a small country in the heart of Europe which has earned worldwide recognition for its long tradition and remarkable results in the area of digital archiving: in 2005 the National Library of the Czech Republic was awarded the first UNESCO/Jikji Memory of the World Prize for its contribution to the preservation and accessibility of our Czech language documentary heritage. For more than ten years now, large national projects have operated in combination with regional, subject-oriented, institutional, and other projects.

These more or less isolated projects have produced millions of digitized pages accompanied by metadata. Now we face the challenge of closer cooperation and coordination of these projects within the framework of the Czech Digital Library.

Newly emerging international standards require changes of standards and workflows developed many years ago. Changes are inevitable, but painful. Existing digital data present not only an important national cultural heritage, but also a burden in changing times.

## Conceptual Content

The plan starts with the international and national structures plus all the valid Czech legislation that underlies those parts that are being - or will have to be - changed to bring the concept to life. Both the strong and weak points of the present situation as well as the risks and opportunities are analyzed, as is the usual practice in planning. A brief explanation of the present situation and formulation of the main goals to be reached by 2010 are accompanied by detailed tables containing sub-goals, necessary actions, funding, etc.

Documents to be preserved and made permanently accessible are divided into two main groups (traditional hardcopy documents and documents in digital form) or four subgroups: traditional documents and three subcategories of digital documents: digitized historical (pre-1801) documents, digitized "new" documents (i.e., produced after 1800), and documents originally created in digital form.

Preservation of and permanent access to traditional and digital documents are becoming less and less independent. Documents printed on acidic paper and/or frequently used have to be digitized to be accessible. The majority of documents published in printed form today are generated by their originally digital form. However, they are delivered to depository libraries mostly only in paper form to wait there for complicated and expensive redigitization in the future. This is an unsustainable situation which needs dramatic changes in legislation, ways of thinking, and cooperation between repositories and publishers. The sooner the waste of allocating insufficient resources is stopped, the more can be saved for the future. From this general observation, valid internationally, let us switch to the situation in the Czech Republic.

The area of preservation of and access to traditional documents is relatively stable in the areas of authorizing legislation, terminology, standards, and the division of responsibility for preservation and access. There are three large depository libraries (which get the complete legal deposit) at the core of the whole system, and they are accompanied by regional and specialized libraries. Legislation has been permanently updated to accommodate newly produced documents, and it does not need to be changed. However, the depository libraries have problems meeting this legal requirement. Good will is not enough in cases of inadequate conditions and funding. The National Library can serve as a good example - its capacity will be totally filled in 2010. There is a plan for a new building, which, if either unrealized or delayed, will result in the breakdown of the whole system from the center. Even when stored in perfect conditions, traditional documents printed on acid paper are seriously endangered. Their digitization is relatively slow, and more and more of them are becoming inaccessible.

As mentioned above, authorizing legislation, terminology, and responsibilities for preservation of and access to traditional documents are relatively stable and satisfactory. However, this is

not the case for digital (both digitized and digitally originated) documents. Legal deposit legislation does not include Internet resources and must be updated. We are lucky that it is so vague in some areas that it allows harvesting.
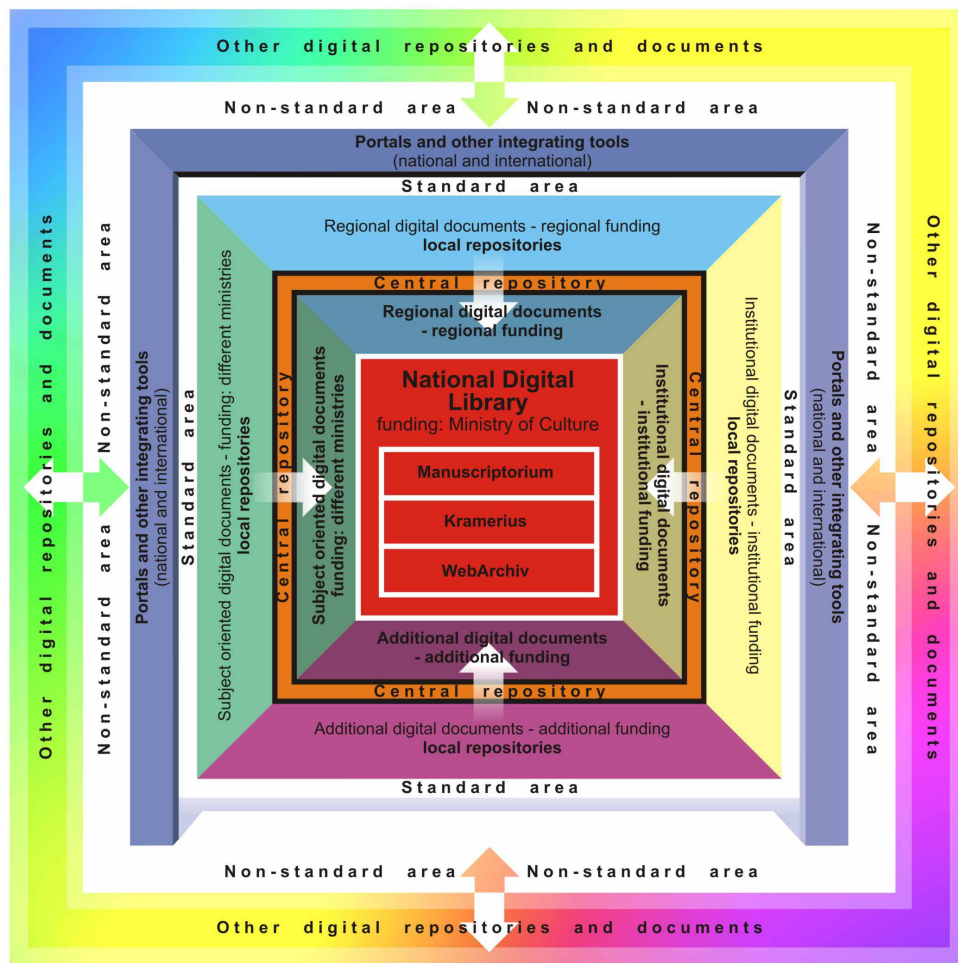
Copyright law (in our region, rather "authorship" law) is a serious burden for access and requires changes as well. Terminology is either nonexistent or just coming into use and requiring agreement. A new division of responsibilities is also needed. Both must be justified. Dividing responsibilities will be probably the easier task, however. Cooperation among libraries in the area of preservation of and permanent access to digital documents can be established relatively easily in countries with a history of good cooperation in building national archives of printed and other traditional documents. This is the case of the Czech Republic. We can easily use the already-functioning structure described above in connection with traditional documents and adapt or develop it for digital documents. This is very practical because nowadays libraries function as hybrid libraries, integrating both traditional and digital documents and building services based on them (and, of course, on resources of other libraries and institutions worldwide).

## The National Digital Library as the Heart of the Czech Digital Library

The figure below illustrates the concept of the Czech Digital Library.

## Czech Digital Library

Let us start in the middle of the diagram. The center, or heart, of the whole system called the National Digital Library contains selected digital documents as the core national cultural heritage. These digital documents intended for a long-term preservation are both digitized or digitally created, and they are produced within three large national projects funded by the Ministry of Culture.

1. "Memoria" http://www.memoria.cz/site_en/index.asp aims to build a virtual research environment for the realm of historical book resources. "Manuscriptorium", the database within Memoria, http://www.memoria.cz/site_en/podmenu_databaze.asp, represents the current state of our achievements.
2. "Kramerius" project http://kramerius.nkp.cz/kramerius/Welcome.do?lang=en focuses on preservation of and accessibility to endangered "modern" periodicals, books, and other documents.
3. "WebArchiv" http://www.webarchiv.cz/index-e.html has two main objectives: first, to catalogue online publications as a part of the national output and to integrate their records into the Czech National Bibliography database; and second, to store these publications in a deposit system so as to ensure long-term access to them.

Documents owned or produced by any Czech library, museum, archive, or other institution can be selected to be part of the National Digital Library. Digitizing, creating metadata, and preserving these selected documents are funded by the Ministry of Culture.

Three different repositories have hosted the separate data of Manuscriptorium, Kramerius, and WebArchiv. From 2006 onward, the digital data of these three subsystems will physically share one large repository.

Institutions with digital data not selected for the National Digital Library will also be invited to deposit their data in the central repository at their request, but additional funding will be required from other ministries (according to different subject areas), regions, or institutions. Other institutions produce digital data but are not interested in depositing their data in the central repository. Digital data stored in local repositories built by such institutions and funded by different ministries, local authorities, individual institutions, and corporations can be integrated under the umbrella of national, or international, portals and other integrating tools, if those agencies adhere to agreed standards.

Of course, very large group of "non-standard" digital documents that do not adhere to agreed standards or newly emerging documents still awaiting standards development should not be omitted. With some limitations, they might be integrated under or linked to portals and other integrating tools. The concept describes different subsystems of the Czech Digital Library together with a system of co-funding from different resources, which is the pre-condition for a viable system.

The first phase of the Czech Digital Library is scheduled for 2006-2010 and will focus mainly on developing standards for digitization (file formats) and metadata, including containers and

.

packaging, preservation metadata, digital rights management and content authentication, storage technologies, and access (including indexing, search and retrieval). Decisions about what to choose for changes and, mainly, the best sequence for making necessary changes (neither too early nor too late) will not be easy; and integrating different existing systems will require some compromises.

## Conclusion

Compared to other European countries, the Czech Republic still belongs among the advanced countries in the area of preservation of and access to both traditional and digital documents. There is a well-established system of preservation and access to traditional documents in our country. Adherence to international standards and bilingual (Czech-English) subject headings make our traditional collections easily available internationally via international integrators, such as WordCat, TEL, or our Uniform Information Gateway portal http://www.jib.cz.

Thanks to large national grant projects, our digitization projects started in the early '90s and achieved archiving of our website in 2000. We have several million digitized pages and about 26 million archived web documents, not bad for a small country. These documents are accessible via the three above-mentioned subsystems, Manuscriptorium, Kramerius and WebArchiv.

We plan to integrate Kramerius and WebArchiv in 2007 under the umbrella of one strong international and tested operational system and to develop Manuscriptorium so that it can cover the special needs of researchers by bringing together fragments of historical documents dispersed in different countries and continents.

Despite the significant progress and proven successes of the past decade, contemplating what the future might bring evokes some sense of fear. Why? The reason is simple. We now have a plan for a new building to house the National Library as the solution for traditional documents and the services based mainly - of course, not only - on them. The new building would also be able to include the expanded concept of a National Digital Library. Both have to be approved and funded by the Czech Government, and they seem to be quite a high priority for the present government. However, we will be holding elections very soon. So, what if they are not approved by a new government, and the necessary sums are not appropriated for realizing these plans? In that case, the Czech cultural heritage might be seriously endangered and perhaps a substantial portion of it irrevocably lost.

## Author Biography

*Bohdana Stoklasova, PhDr., is the Acting Director of Library Collections and Services at the National Library of the Czech Republic. She is responsible for coordination of different national programs connected with her area of responsibility, including the national Uniform Information Gateway portal and the Czech Digital Library.*