

## **ANNEX 1: Collaboration between partners**

### ***National Library of the Czech Republic***

Due to Culture2000 Project we gained new, self-contained selection criteria which will definitely help us in our future work. Definition of criteria is mainly important for technical issues of the web archiving. It enables us to set the automatic harvesting optimally and thereby to create more consistent archive of the Czech web resources.

We started to build the Czech web archive six years ago, so unlike our partners we had at our disposal some technical and technological facilities having been long-run tested. It was interesting for us to assist the partners to identify their national web space and also to help them to set suitable technical parameters. It was very useful and interesting as well to compare our tools, i.e. our crawler applied for harvesting with our partners' harvesters.

As a project leader we had to solve many problems – personnel, technical, organizational, administrative which was sometimes very demanding because we had no experience in coordinating any EU project before. Nevertheless, our partners always tried to help us and had very matter-of-fact approach to problems. We are convinced that this collegial co-operation helped significantly to reach the main goals of the project.

In the frame of the project co-operation there arose new relationships which, we hope, will continue in future.

### ***National Library of Estonia***

Before the project Web Cultural Heritage the National Library of Estonia had been engaged only with downloading separated files from the Internet, using free program HTTrack for this. During the project we started to download websites, which was the next step for our library on archiving whole domain .ee in the future.

We also experimented with the program Heritrix. In the result of testing the National Library of Estonia will start to use Heritrix for archiving Estonian websites.

Getting to know about the experiences of our partners we detailed our selection criteria of web archiving during the project.

Our library got new cooperation partners and colleagues from the other European libraries and the exchanging experiences was hopefully useful for everybody.

### ***National and University Library in Ljubljana, Slovenia***

The project was definitely too short. We believe that better results and more dissemination would be achieved with at least additional 6 months of work. For the National and University Library (NUK) it was very important the moment of starting the project, as we were just implementing a web archive and defining the selection criteria. The comparative analysis and different experiences from partners' institutions gave us broader dimensions in this regard.

The project discussions also helped us in the definition of technical functionalities of the crawler. NUK crawler was developed by The Jozef Stefan Institute (IJS). The specifications of the Brno University Library and their requirements were very useful in the adjustment of the tool to our needs. We discovered some errors and problems in our crawler that the researchers from JSI had to correct. For example, focused crawling was not completely working, and some technical elements in some domains were blocking the system. These adjustments and corrections produced also big delays from our side in the project and we would like to have had more time for testing and evaluating the crawler. The researchers of IJS are finishing an access module (indexing and searching) to the web archive based on Nutch technology. This module was not possible to complete till the end of the project to compare it with Czech and Estonian solutions.

The cooperation in the project was very useful also in the sense of creating a network of partners dedicated to web archiving as well as for the exchange of experiences in the future. Brno University Library was the most experienced partner in the integration of web archiving in the library system. It was very helpful for us in the communication with the computer scientists researchers from IJS.

### ***University Library in Bratislava, Slovakia***

We consider the project co-operation as inspirative, educative and very helpful. University library in Bratislava (ULB) and also Slovakia had no web-archiving activity before the Culture2000 project. Co-operation with our project partners enabled us to start a pilot harvest of our web. ULB established a web-archiving infrastructure, with support of our Czech colleagues. We already made the first pilot harvest of electronic serials registered by the Slovak ISSN agency.

There are about 92.000 sites under the Slovak top level domain .sk, registered by almost 47.000 users. Our first results show, that the selected set of websites - 191 born digital e-journals - look as follows. Average web (e-journal site) has 604,23 MB and 24.496,44 documents, the average object has 25 KB.

These results, quantitative and qualitative analysis of the testing harvest will help us to identify the Slovak web, to set the optimal technical and personal capacity for web-archiving. We have to increase our storage capacity, to make conclusions with publishers, to organize the workflow. Thanks to the Culture2000 project, we can use some know-how from our partners and adjust the selection criteria recommendations for our purposes.