2nd report on the actions taken to give effect to recommendations as formulated in the 2003 October UNESCO General Conference concerning the promotion and use of multilingualism and universal access to cyberspace
– Hungary –

1 Introduction

Focusing on the developments and activities in Hungary in recent years with respect to the advancement of multilingualism and universal access to information in cyberspace the present report addresses the relevant issues in the order as laid out in the recommendation.

2 Development of multilingual content and systems

Due to the unfavourable economic environment of recent years national programs have seen a considerable downturn with respect to the advancement of the creation and processing of educational, cultural and scientific content in digital form as well as facilitating access to it. Nonetheless, some significant initiatives have been launched to support the development of (multilingual) digital content and innovative communication technologies.

• *National Research Infrastructure Survey and Roadmap*. The aim of the project is to survey state-of-the art national research infrastructures and to develop a coherent policy for the implementation of these infrastructures. Language resources and applications developed in Hungary in recent years have been assigned the status of *strategic research infrastructure* within this project.

To single out just a few prominent resources and applications a list of recent developments are presented below.

– *Hungarian EuroWordnet*. The Hungarian version of the multilingual database with wordnets for several European languages representing a unique language-internal system of lexicalizations. The database can be used to develop intelligent linguistic systems, for monolingual and cross-lingual information retrieval, machine translation etc.

– *Complex Uralic Linguistic Database*. Annotated databases and language processing tools for endangered Uralic languages.

– *Ob-Ugric Linguistic Resources*. Modern descriptive online resource on Ob-Ugric languages, including corpora, e-dictionary, e-grammar and language processing tools.

– *Language processing tools for Hungarian*. In spite of the significant decrease in available project funding further natural language processing applications have been developed in numerous Hungarian research centers, now more and more focusing on higher order language processing tasks, such as deep parsing and semantic analysis.

• *National Technology Platforms.* Supported by the National Office for Research and Technology, a new network of technology platforms has been established, a number of which focusing on domains with direct relevance to the proposed recommendations, such as:

– *Language and Speech Technology Platform*, which defines some relevant strategic priorities as the following:

1. Setting up and maintaining a national research infrastructure in the field of language and speech technology
2. Managing, storing and processing of linguistic information
3. Preservation of our linguistic cultural heritage in the digital age
4. Natural human-machine communication
5. Multilingualism, overcoming language barriers

– *Creative Industries Technology Platform*, which focuses its activities (among others) on the promotion of the internationally accredited Hungarian educational achievements by value-added digitization and the development of their networked training systems; the digital dissemination technologies of books; development of networked learning environments with special interests on disabled and elderly people; creating a copyright-friendly electronic environment; digital content creation.

– *NESSI-Hungary Software and Services National Technology Platform*, which aims to elaborate development and research strategies in information service architecture and software infrastructure to serve a knowledge-based society.

To facilitate digital publication, to ensure universal access to cultural and scientific content in digital form, several public repositories have been significantly extended in recent years:

• *National Digital Data Archives*.[1] A central register of the cultural and public contents produced in the Hungarian language and available digitally on the Internet. It is designed to facilitate the availability and the reuse of public data, and is permanently extended. At present about 500.000 documents are available through its service from virtual libraries and museums, and digital archives.

• *National Audio-Visual Archive*.[2] A deposit archive of the Hungarian national broadcasters, collecting audiovisual content. Its scope of collection comprises the programmes of the national terrestrial televisions and radios, produced in or relating to Hungary. In addition, it also accepts locally broadcast or any other audiovisual content for processing or storing, facilitating the preservation of audiovisual content as part of the cultural heritage.

• *Digital Literary Academy*.[3] The long term support and maintenance of this wide coverage database of the works of contemporary Hungarian authors has been established and the permanent extension of the collection has been ensured.

---

[1] http://www.nda.hu

[2] http://www.nava.hu

[3] http://www.irodalmiakademia.hu

- *Bibliotheca Hungarica Internetiana*. The 400 works of outstanding authors ranging from old Hungarian literature through the greatest classics of the early 20th century have been incorporated into the *Hungarian Electronic Library*.

- *Hungarian Electronic Library*.[4] Still being one of the most popular and most significant text-archives of the Hungarian webspace, this online collection of resources concerning Hungary or the Central European region is now maintained by the Hungarian National Széchényi Library.

- *Digital Library of the National Széchényi Library*[5] (abbr.: NSZL-DL). This digital collection can be accessed anytime from anywhere via the Internet. The NSZL-DL is a huge virtual library, which can be searched or browsed in the same way as a physical library, with the difference that the chosen reading can be read immediately with one click. The National Széchényi Library is a legal deposit for books published in Hungary. According to the law on legal deposits, publishers are obligated to send 6 copies of each published book to the National Széchényi Library. This system is working already virtually, as the basis of the Digital Library. Through this Submission system publishers are able to send the documents without delay and postage. After integrating into the searching and browsing system, the received document is ready to be accessed by the readers.

In a recent project supported by the National Development Agency, the *Sulinet*[6] portal, providing a complex digital knowledge-base of education materials, has been completely redesigned and updated, with extensive new content uploaded.

To provide shorter or longer term funding for the development of info-communication technologies, in particular, natural language processing applications and language technologies, there is still a great need for national programs. In recent years only a few of such projects have been able to launch with the (decreasing) support of the National Office for Research and Technology or the National Development Agency. Some of the prominent developments in the field of intelligent linguistic systems and network services are the following:

- MASZEKER.[7] Model-based semantic search system being developed for English and Hungarian, based on a novel approach to internet search engines.

- TEXTREND.[8] An free integrated toolkit and platform for the in-depth analysis of dynamic information on the web for science or business.

3 Facilitating access to networks and services

---

[4] http://mek.oszk.hu

[5] http://oszkdk.oszk.hu

[6] "school-net" – http://www.sulinet.hu

[7] http://www.maszeker.hu

[8] http://www.textrend.hu

Funded by several national and governmental projects, the *National Information Infrastructure Development Institute*[9] is responsible for serving the framework for the development and operation of the research network in Hungary. A significant recent development of the institute is the new generation broadband network (HBONE+), which establishes a direct connection to the pan-European data network, GÉANT, dedicated to the research and education community, and makes its services available for the Hungarian community. Another related initiative, WEB2GRID[10] under the coordination of the Computer and Automation Research Institute of the Hungarian Academy of Sciences has been launched to support new generation community WEB services and applications by secure business GRID platform.

With respect to the promotion of internet services, the Strategic Research agenda of the NESSI-Hungary Platform claims: "The rate of ICT R&D expenses in Hungary is low even in comparison with the European average. However, some ICT indices, such as the broadband Internet penetration and even more its growth rate now show a much better picture. This is a result of the relative underdevelopment, since the omission of the technological steps favours the spreading of the most modern technologies. Thus there is a possibility for ICT technologies to play an ever more comprehensive social role, but its conditions must be created." It is to be noted that telecommunications and Internet costs are still not affordable for many, making the use of services restricted to only an (albeit rapidly increasing) minority of the population, the general public.

4 Development of public domain content

To enable universal online access to public and government-held records at a national level, and to comply with Act XC. of 2005 on the Freedom of Electronic Information, the services of the general web portal for public administration[11] has been significantly extended in recent years. The development of similar sites bringing together the widest range of public service information and services online is now a de facto standard at the level of local councils, offering on-line access to relevant information and services. In line with the above law, all authorities performing a public role are liable to grant access to their public data specified by the law through their websites, free of charge and registration.

Within the framework of the *Public Sector Information Program*, a central electronic register[12] keeps track of the descriptive data of the available information from organizations, and a related uniform public data search system has been linked to this resource[13], with additional links to collections of official data resources of public interest.

---

[9] http://www.niif.hu

[10] http://www.sztaki.hu/search/projects/project_information/?uid=00274

[11] www.magyarorszag.hu

[12] www.kozadat.hu

[13] www.kozadattar.hu

5 Reaffirming the equitable balance between the interests of rights-holders and the public interest

For establishing easily-available shared resources for research communities and also for the general public it is essential that appropriate copyright legislation be developed.

6 Summary

This report has discussed in brief the initiatives taken to give effect to the norms and principles set forth in the recommendation with special emphasis on the development of language resources in digital form as well as those in computer aided language processing. It is necessary to provide national support to programs focusing on the objectives outlined as strategic priorities in Section 2, due the relatively small number of Hungarian speakers without sufficient market demand, to allow for the efficient use of resources and ICT technology and to facilitate a multilingual and multicultural education with enabling lower thresholds to multicultural and multilingual content accessible for the general society.