

Hungary

Ivan Rónai
Ministry of Cultural Heritage

A. Users and content

The inventory of the Hungarian cultural heritage shows that we have to deal with a huge quantity. The number of units to be digitised was determined by taking into consideration the needs of the user community. Progress made in 2005 in the field of cultural digitisation was centred around the major memory institutions namely the national library (OSzK) and large municipal libraries, the national archives, the main state museums, the national film archives, and

also the newly established electronic collections such as the Neumann Digital Library, the National Digital Data Archive (NDA) and the National Audiovisual Archive (NAVA). The Office of the Protection of Cultural Heritage (KÖH), which is responsible for the preservation of monuments and archaeological sites, has performed a notable step forward digitising source documents and images of the immovable heritage and also the data related to archaeological research. Hereby, a brief description is given about the progress in the main cultural fields.

Progress in the area of immovable heritage

	Inventory units	Units to be digitised	Units digitised
Immovable heritage	11 000 monuments 50 000 archaeological sites 915 000 source documents	915 000 source documents	76 948 source documents = 8% of total = 8 MB of 2,1 TB

In 2005 the percentage of the digitised source documents related to monuments (such as photos, blueprints, manuscripts and other records) has doubled and is now very close to the 10 %, which still leaves enough work to be done. The digitisation of 5,000 archaeological sites has been performed for the topographical database of the archaeological registry. This constitutes the tenth of the whole to be entered into the database.

Progress in digitisation of library materials

Although, to our date only a small part of the total has been transformed into digital format, in 2005 a significant step forward was taken in order to make these cultural assets available on the Internet and/or in electronic readable format.

	Inventory units	Units to be digitised
Library	111 000 000 documents	3 520 000 documents + 2 800 000 pages + 330 000 hours

By type of documents:

Type	Quantity	Quality
Antiquity	500	pieces
Small print	1 600 000	pieces
Journal volumes (textual printed doc.)	25 000	pieces
Photograph	1 500 000	pieces
Sound recording	250 000	hours
School yearbook	20 000	pieces
Cartographic documents	100 000	pieces
Manuscript	60 000	pieces
Manuscript	1 500 000	pages
Codes, incunabula	300	pieces
Notes	50 000	pieces
Book (textual printed doc.)	100 000	pieces
Archival library documents	1 300 000	pages
Engravings, arts documents	20 000	pieces
Theatre scenarios (handwritten, typewritten, printed)	40 000	pieces
Blueprint	4 000	pages
Video documents	80 000	hours
Other documents	5 000	pieces

GOOD PRACTICE

Digital Core Library of Hungarology
(Hungarológiai Alapkönyvtár)

The transformation of about 3,800, mostly voluminous works into electronic format is the largest digitisation enterprise in Hungary at the moment.

Hungarology, or Hungarian studies, covers interdisciplinary research dealing with the Hungarian people, its history, language, culture, national civilisation in the past and present, as well as its place and role in the universal civilisation of mankind.

As such, in addition to the disciplines listed, it focuses on history, language and literature, fiction, cultural history, ethnological research and geography.

The first serious initiative in this area was the publication of a book list titled "A core library of Hungarian studies" in 1986. The next publication to follow was the "List of reference works" compiled by the National Széchényi Library Centre for Library Science and Methodology (CLSM) in 1993. It has focused on handbooks and gave an overview of the materials in a variety of disciplines. This list was

therefore shorter and more modest than its predecessor.

The project to enable remote access to the original works in the list is coordinated by the National Széchényi Library; the digital versions are made available in the Hungarian Electronic Library <<http://mek.oszk.hu/hungalap/indexeng.html>>. The work started in 2004. At first, the most important academic handbooks – the materials of the 1986 and 1993 book lists, plus the additions from 2003 – get published on the Internet.

Number of works to be digitised	~ 3 800
Number of digitised works	49
Number of digitised pages	~ 58 541
Number of digitised characters:	~ 149 623 000
Number of digitised images	~ 12 000

Use of standards

HTML, MP3, PDF

Metadata: HUNMARC, USMARC, XML, Dublin Core

Digitisation of the journal «Vasárnapi Újság» 1854-1921

In the first phase of the project the issues of 11 volumes between 1861 and 1871 were transformed into digital. Altogether 7,000 full text pages in pdf format are available and searchable. The bigger part that is 4,000 pages were digitised in 2005.

Digital Literary Academy

For the past 7 years of the project the works of 63 contemporary Hungarian writers and poets have been digitised and made accessible on the Internet. In 2005 the number of digitised works increased by 3 thousand and totalled 35,000. As storage format SGLM standard is used and data are appended with metadata in qualified DC. The content and the services are of high quality, they follow W3C recommendations.

<<http://www.irodalmiakademia.hu>>

Progress in digitisation of archives material

	Inventory units	Units to be digitised	Units digitised
Archives	3 500 000 000 pages	40 000 000 units	1 500 000

Archival documents of the Hungarian government (from 1867 to 1944), Foreign Ministry, the Hungarian Socialist Workers' Party (1956-1989) and the Hungarian Workers Party (1948-1956) were transformed into digital images using TIF or JPG format. In addition, 1923-1989 issues of the periodical "Archives Bulletin" were digitised.

Hungary's medieval digital archival documents 1,0
The digitisation of approximately 60 volumes of medieval codes and certificates was completed and published in Hungarian, German and English languages.

GOOD PRACTICE

Progress in digitisation of museum material

	Inventory units	Units to be digitised
Museum	59 000 000 objects	14 800 000

Breakdown of museum holdings to be digitised by type (Approximately 25% of the whole inventory)

Type	Quality
Collection objects	3 200 000 pieces
Original documentation of source value	1 275 000 pieces
Photo negative	1 300 000 pieces
Slides (diapositive)	190 000 pieces
Data archive units	450 000 pieces
Mass registry units	6 000 000 pieces
Estimated non-inventorised quantity	2 400 000 pieces
Total:	14 800 000 pieces

GOOD PRACTICE

The Open-air Museum of Ethnography
It has completed the digitisation of its study collection. It contains 2,300 objects with 3-5 images per object. The archive of photo negatives of the Hungarian Folks Architecture was transferred into digital format comprising approximately 70,000 images and the pertaining documentation. In 2005 the digital version of the Slides collection in this archive grew to more than 18000 registry cards with almost 7,000 images. In the collection of the Hungarian Folks Architecture 17,663 registry cards and 6070 scanned images were digitised. In 2005 the digital collection of folk art objects considerably grew in number and when completed, it will contain 44,861 objects and more than 20,000 images.
<<http://www.skanzen.hu>>

Progress in digitisation of the audiovisual heritage

	Inventory units	Units to be digitised	Units digitised
Audiovisual	20 000 films	5 000 films	100 feature films
	165 000 hours tv	165 000 hours tv	+590 newsreel 20+ 750 hours tv
	662 000 radio tapes	662 000 radio tapes	1000 hours of sound
	44 000 000 pages of documentation	44 000 000 pages	1000 pages

Film heritage

In 2005 the Hungarian National Film Archives completed the digitisation of 100 feature films from the 20th century. The selection criteria were manifold: importance from film history aspects; preservation of films from the 50s in poor condition and also colour films from the 70s; the inclusion of motion picture works representing all the major trends in film creation. In addition, 590 newsreel documents from the years between 1931 and 1944 were digitised. This selection of the digitised film heritage is an outstanding e-learning tool for literature and history teaching in secondary education.

GOOD PRACTICE

In 2005 a small but very important segment of the Hungarian audiovisual heritage was also digitised: The chronology of the Hungarian revolution and freedom fight of 1956. The database running on an open source OS contains the happenings of the historic days (from 22 October to 13 November 1956) minute by minute taken from the following material:

- News agencies' reports
- The monitoring of the broadcasts of the Kossuth Radio of Budapest
- The monitoring of the broadcasts of the Radio Free Europe of Munich (in both cases sound recordings and their textual transcriptions are retrievable)
- Digitised versions of major contemporary Hungarian newspapers
- The fullest digitised collection of film documents of the revolution
- Memories and memoirs of "storytellers" of the epoch (radio, newsreel reporters and journalists)
- Indices of names, sources, literature and pseudonym lexicon.

The digital chronology of the Hungarian revolution and freedom fight contains 1000 pages of newspaper text, 1,000 hours of sound plus their textual adaptation, 1,200 minutes of digitised film documents.

Media observer of the National Széchényi Library
From 1 January 2001 key news programs of 5

nationwide broadcasting televisions (m1, m2, Duna TV, TV2, RTL Klub) are recorded digitally (1Mbit/sec) and processed in order to make them available for research and long-term preservation. The database contains 45,000 minutes of program plus the searchable textfile version of the news broadcast.

The National Audiovisual Archive

NAVA (National Audiovisual Archive) constitutes the legal deposit archive of the Hungarian national broadcasters. According to the Act on NAVA the legal deposit is compulsory from 1 January 2006. Its scope of collection comprises the programmes of the national terrestrial televisions and radios, produced in Hungary. NAVA provides on-line access to its collection within the framework regulated by law. In addition, it also helps in processing local broadcasters' archives.

NAVA started recording and archiving broadcasters' material in 2005, and by the end of the year its database contained more than 13,000 records. The database can be searched by keywords of DC identifiers, time of broadcasting and other criteria. As a result of the search the user gets a description of the program and a series of snapshots from the program. According to IPR regulations only users of a dedicated network of educational and library institutions are permitted full access to the records.

<<http://www.nava.hu>>

Hereby see an illustration of a search result:



B. Technologies for digitisation

In many projects – especially in those of best practices – the latest equipment is implemented like 6,000x6,000 dpi photo scanner, A0 / 400 dpi document scanners. However in many cases digitisation activities lack the systematic approach. Although there is a tendency to use descriptions and metadata standards, in many projects digitisation is limited to scanning images. The use of new technologies also includes the implementation of special image compression methods, map displaying programs and new text converter programs for more user friendly text handling. It must be mentioned here that the National Digital Data Archive (NDA: <http://www.nda.hu>) aims to provide a content infrastructure for digital cultural content of memory institutions.

The operating model of the NDA is based on the Open Archive Initiative (OAI). It offers a new way of publication based on the separation of data and service provider functions. The data providers (memory institutions) hold digital contents and documents, as well as produce and maintain the information describing the contents. According to the OAI principle, the content remains in the management of data providers, but the level of descriptive data has to be opened for free-access use by the service providers. Due to this latter element it is open archiving model. Its operation is essential, since current service demand of a wide variety of communities can be ensured only through the free-access use of descriptive data (metainformation). The technical operation of NDA is based OAI-PMH 2.0 (Protocol for Metadata Harvesting) protocol. The structure of the OAI-PMH data-exchange is based on the pattern of Dublin Core (DC) that is widely known and accepted worldwide, too.

C. Sustainability of content

Public memory institutions usually do not have significant funds in their budget for digitisation. They keep an open eye on grants and other state fundings to carry out their digitisation projects.

Public Private Partnership is rarely exercised. A remarkable example though is Arcanum – a private enterprise –, which is performing the greater part of digital text conversions in Hungary. Many of the above

mentioned digitisation projects are realised in cooperation with Arcanum. A good example of PPP is that Arcanum made the digitisation of old maps in exchange for the publishing rights.
<<http://www.arcanum.hu>>

D. Digital preservation

Long term preservation is not yet evidence in most of the memory institutions. It is acknowledged that measures for long term preservation are a necessity, but with the exception of the National Library, not much have been done so far.

For security considerations, institutions make and store duplicates of the digitised material, and for example the Open-air Museum of Ethnography saves all its databases on three different computers.

E. Concluding remarks

The year 2005 did not bring about dramatic changes in the field of cultural digitisation in Hungary. As digitisation is a time and labour consuming endeavour, it can be stated that we need to take more efforts to speed up our work. The participation in the MINERVA Plus project helped our memory institutions to learn good practices in cultural digitisation and by implementing the MINERVA quality principles they enhance their services to access digital contents. We are still in need of good business models to ensure the funds of mass digitisation, we still have to find the balanced solution in IPR issues, and we must strengthen the cooperation and coordination at national and European level in order to avoid waste of resources and repetition of similar mistakes.

