

Hungary

Iván Rónai
Ministry of Education and Culture

A. Users and content

What are the main materials/categories of objects being digitised?

In museums

- objects of the main museum holdings
- slides and photographic films
- objects of virtual exhibitions

In libraries

- codes, incunabula;
- journal volumes,
- books pertaining to “the core library of Hungarian studies”

In archives

- cadastral maps, diplomatic archives;
- film heritage;
- audiovisual archive – broadcast of nationwide tv and radio programmes;

In sites and monuments protection (archaeology, architecture)

- monuments, protected objects;
- topographies of archaeological sites;
- source documents.

How is the digitised CH content being distributed?

- national inventories: NDA (National Digital Data Archive) is the most important service to foster the visibility of digital cultural content
- portals, internet services,
- other – see the good examples below:

Babelmatrix

When two people from different European countries meet, they often use English as their common language. Literature can also play a great part in learning about each other's cultures. It is transmitted from one language to the other, thanks to the efforts of literary translators. Of course Babel Web Anthology offers only a glimpse into the huge bulk of literary translations. Their multiplicity can easily be handled by the possibilities of the Internet.

The inclusion of texts both in the original and in translation allows for the promotion of one's national literature as well as first taking a glimpse into some other nation's literature through excerpts and then searching for the actual books.

We offer short excerpts only as Babel Web Anthology is intended for reading from display or printing a few pages.

Babelmatrix.org has been initiated by Typotex Ltd. and realised with the contribution of Russicon Company. The Project started with the support gained at the e-content application (ITP-6/4) of the Hungarian Ministry for Informatics and Telecommunication. In 2005 the project was supported by GVOP (4.2.1 2004-07-0015/4.0).

The portal awarded by the Hungarian eFestival.

Babelmatrix.org offers literary works or excerpts thereof under clean copyright conditions only. Using the anthology for commercial publications is allowed exclusively on previous permission. In some cases, although we have made significant efforts to find copyright owners of certain publications, we did not succeed. So if you are a copyright owner of one of excerpts displayed in babelmatrix, and we did not contact you yet, please contact us.

<copyright@babelmatrix.org>.

Hungarian language

The Hungarian language is presented by 462 works of 131 authors

There are the following translations into Hungarian:

Czech	language:	<u>74</u>	works of	<u>52</u>	authors
Dutch	language:	<u>22</u>	works of	<u>11</u>	authors
English	language:	<u>59</u>	works of	<u>26</u>	authors
Finnish	language:	<u>1</u>	works of	<u>1</u>	authors
French	language:	<u>59</u>	works of	<u>27</u>	authors
German	language:	<u>33</u>	works of	<u>20</u>	authors
Italian	language:	<u>9</u>	works of	<u>7</u>	authors
Polish	language:	<u>53</u>	works of	<u>40</u>	authors
Portuguese	language:	<u>11</u>	works of	<u>7</u>	authors
Russian	language:	<u>151</u>	works of	<u>59</u>	authors
Slovak	language:	<u>46</u>	works of	<u>21</u>	authors
Spanish	language:	<u>10</u>	works of	<u>10</u>	authors

Photo collection of Hungarian film history

The collection <http://www.mafilm.hu/1024/fotoarc_en.html> contains photographs taken at the shootings of Hungarian feature films. The database comprises the textual metadata of 920 films and 17500 digitised work photos of 700 films. It can be searched by many aspects:

- title or work title of the film
- protagonists (actors, actresses etc.) of the film
- producer
- director
- year of production
- prizes awarded to the films
- still photographer
- and many other contributors of the film.

As a result of the search, the digital photos can be displayed on the screen as watermarked low resolution images. A business model has been worked out for those who want to use the high resolution digital photos for commercial purposes.

Are there any eLearning or creative industry uses for the digitised CH material?

A good example for eLearning tools is The Momentous Days, an educational assistance material, which is related to several areas of science and knowledge, so it can be used in teaching many subjects.

The Momentous Days is built around the calendar system and provides rich materials presenting the traditional and new jubilees, as well as the feasts already forgotten. The core of the collection is a list of festive days, which consists of more than 1000 momentous days from the European and mainly Hungarian tradition. Besides National holidays and Hungarian-related jubilees, there are also spiritual feasts, World Days and International Days on the homepage.

An Additional value is that the collection is an open system so it can be further extended with new jubilees, feasts, texts and pictures. Libraries, museums, and schools are invited to develop the "Momentous Days" together according to the needs.

Are there any special issues that deal with the general accessibility of the digitised cultural content?

The copyright issue – enabling accessibility of digitised copyrighted material on dedicated terminals for educational and scientific purposes.

According to the Act No. LXXVI of 1999 on Copyright the public libraries and public collections may freely publish their digital collection in their closed network. In Section 38 Paragraph (5) of the Act No. LXXVI of 1999 on Copyright, Section 66 of the Act No. CII of 2003 on the amendment of certain industrial right protection and copyright acts completed the present paragraph with «for lack of any differing agreement on use, works belonging to collections of libraries providing public services, institutions serving school education, museum institutions, public record offices, audio and picture archives shall be freely presented to individuals of the public, on the screens of computer terminals installed to this end on the premises of such institutions, for scientific research and individual learning purposes; and in its interest – on ways and conditions stipulated in a special legal regulation – they shall be transmitted to the aforementioned members of the public, including provision of accessibility, supposing that such use neither indirectly serves purposes of income earning or income growth.»

B. Technologies for digitisation

What common standards are most commonly used in digitisation?

National Audiovisual Archive (NAVA) is a completely digital system, even the media storage is carried out in a data format.

While choosing the formats, our concept is that the accessible quality should follow the broadcasting quality provided by the given broadcasting company and it should provide extended services at the bandwidth largely offered by internet service providers. Storing at a reusable quality is not a primary objective.

NAVA's storage quality (open standard): MPEG2, MP@ML (720x576, 25 fps), CBR 8MBps.

NAVA's service quality (free player): RealMedia, 384x288, 25fps, 512 kbps, RV10 + RA8.

NAVA's servicing interface, practically just as its internal structure, entirely complies with international (where applicable, open) standards, recommendations and initiatives, thus providing an opportunity for easy connection to other archives.

The most important connecting points are:

- Motion Pictures Expert Group (MPEG)
- Dublin Core Metadata Initiative (DCMI)
- European Broadcasting Union (EBU)
- Open Archives Initiative (OAI).

What technologies and tools are most commonly used in digitisation?

In projects of best practices the latest technology is implemented like 6,000x6,000 dpi photo scanner, A0 / 400 dpi document scanners. 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.

According to our latest experiments, digitization of books of the National Library can be made by 40 book scanners in two shifts, 16 hours a day. Thus in average 3048 documents can be processed per day. The result will be good quality TIF files.

The 40 scanners would work on a common storage

of high capacity and quality allowing for some buffering, as well.

The operating model of National Digital Data Archive (NDA) is based on the Open Archive Initiative (OAI). 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.

Concerning the NAVA the following requirements arise from the border conditions (based on the programme structure and the preliminary analysis of consumers' reception facilities):

- recording 5 input video channels
- storing video recordings of 25,000 hours a year (38 hours a day)
- generating and storing descriptive data of 100,000 programme items a year.

How is interoperability ensured?

On the one hand, interoperability is ensured by using universal identifiers. pages you can register, manipulate and resolve URN:NBN identifiers (URN — Universal Resource Number, NBN — National Bibliography Number) which supports the long-term identification of the documents on the web. On the other hand, it is required from applicants for state grants on digitisation to create metadata compliant with DC and qDC schemes.

C. Sustainability of content

What funding models and funding sources are most commonly used in digitisation?

Most common way of funding digitisation projects is the application for state grants. Various ministries and the National Cultural Fund issue calls for proposals. In the coming years an important funding source for cultural and scientific digitisation will be the Operational Programme of the National Development Plan. The criteria of sustainability of content creation will be emphasised in the calls for proposals.

Have you any examples of cost-reduction?

A significant cost reduction can be achieved by exporting the documents into countries or regions where digitisation labour costs are much lower. However, this solution may have risks of damage to the library holdings.

Have you any examples of national competence centres in digitisation of cultural content?

In 2007 two competence centres are planned to be set up: one in the National Library functioning as a national knowledge centre for library digitisation, the second in the Neumann Digital Library concentrating on digitisation of audiovisual (broadcast media), museum and archival material.

In 2006 the Open-air Museum of Ethnography has completed the digitisation of 3 collections containing 60 thousand cultural objects, 75 thousand slides, and 17

thousand pages of building documentation of the Hungarian Folks Architecture collection. In addition, the documentary films presenting the permanent exhibitions of the museum have been digitised.

D. Digital preservation

Are there national policies for digital preservation?

The National Library is responsible for collecting and preserving digital cultural objects with the exception of TV and radio broadcasts, archival material, digitised holdings of museums, and digitised units of the immovable heritage (monuments, archaeological sites, and source documents). Born digital documents are subject to legal deposit and are stored and preserved by the National Library. The long-term preservation issue is presently under consideration.

E. Monitoring progress

Input/output indicators

Input	MEUR / Posts	Comment
Funding allocated to support cultural digitisation - government	5,4	estimation for 2006
Funding allocated from within existing institutional budgets*	1,3	estimation for 2006
Size of work-force (full-time equivalent posts)	175	estimation for 2006

*Combined total, including personnel costs.

Output	Number of collections / items	Comment
Number of digitised collections		all numbers are estimations
- museums	60	Partly digitised
- libraries	90	Partly digitised
- archives	15	Partly digitised
Number of digitised collections recorded in a national inventory		
- museums	35	
- libraries	23	
- archives (audiovisual included)	8	
Number of digitised items per collection type		all numbers are estimations
- object entries (m/i)		
- text documents (m/i)	1,2 million in catalogues 14 thousand digitised	
- images (m/i)		
- audio (m/i)		
- moving images (m/i)	100 films digitised 920 films metadata + 17500 werk photos digitised	
tv and radio broadcast programmes	248 thousand items catalogued and digitised	

The following collections are recorded in a national inventory = accessible at least on collection level

through the national inventory (web portal, web service) NDA:

Name	Description	No. of digital objects
NAVA	Digital archive of TV-programs	248 000
MATARKA	Database of CC of Hungarian periodicals - catalogue of articles published by on-line periodicals	52 285
Skanzen (Szentendre)	Open-air Museum of Ethnography - digitized slides collection of Folks Architecture	45 312
Kurt Lewin Alapítvány	Kurt Lewin Foundation - articles in online RomaPage and SzocHáló (social net)	18 005
Neumann-ház	Neumann Digital Library - digitised works of fiction and professional literature	17 315
56-os Intézet	Public Foundation of the History of the 1956 Hungarian Revolution - documents of the 1956 Revolution	8 487
Váli Dezső honlapja	Váli Dezső oeuvres - paintings and other works of art	8 337
Postamúzeum	Museum of Post and Telecommunication - digitised photo collection	6 195
HUNG-ART	MEKE - fine arts in Hungary, a virtual museum	5 167
OSZK (MEK)	National Széchényi Library - Hungarian Electronic Library - digitised texts of fiction and professional literature	4 734
BMC	Budapest Music Center Hungarian Music Information Centre - Hungarian music depositary	4 501
KÖH-műtárgyak	Office for the Protection of Cultural Heritage - cultural objects under protection	3 725

