

# Multimedia Oral History Database <sup>1</sup>

This Internet-accessible multimedia oral history database (<http://server2001.rev.hu/oha/index.html>) forms a slice, or cross-section, of the database of contemporary history held by the 1956 Institute in Budapest. Its basis is the Oral History Archive (OHA) established by some of the founders of the institute decades ago.

There are about a thousand life interviews, divisible into three main groups. About 500 interviews were made with participants in the 1956 Hungarian Revolution. Many of the rest were done with their children. Others were life interviews made under a leadership research programme in 1981–5, with those thought likely to have a career ahead of them in the communist party or state hierarchy.

The interviews were tape-recorded. Transcripts were made from these recordings, which up to the end of the 1980s meant that they were typed. In the 1990s, the transcription was made using computer word processing, so that the texts have survived in digital form as well.

In 1991, to assist orientation among the interviews, we developed a database using 3–5-page abstracts whose information could be searched and recovered in detail. The purpose of this database was to record the interview content as accurately as possible. It was not possible to archive the

*by Zoltán Lux \**

full texts or the sound materials in digital form because of the memory constraints at that time.

In 2003, the institute's successful application for competitive funding under the eVilág project of the Ministry of Informatics and Telecommunications allowed a sizeable proportion of the interview materials to be digitized. This funding will allow the Institute to digitize the texts of interviews available only on paper as well as old sound materials. The project involved more than archiving, as the public part of the database is being placed in the public domain via the Internet.

Before the project began, we were developing a database handler into which we could transfer the old oral history database. This Oracle-based database can now receive the full interview texts and the sound documents. This database not only assists with making the content and technical data of the interviews searchable, it is becoming increasingly suited to fulfilling the complete archiving function and digital storage of all the related text and audio-visual documents.

The oral histories comprise only part of our database of

The screenshot shows a web browser window with the URL [http://server2001.rev.hu/oha/index\\_eng.html](http://server2001.rev.hu/oha/index_eng.html). The page features a header with the text "The Institute for the History of the 1956 Hungarian Revolution" and "private history 1956 AND THE KÁDÁR PERIOD". A sidebar on the left contains navigation links: "magyar", "english", "PRIVATE HISTORY, 1956 AND THE KÁDÁR PERIOD", "PREFACE", "GLOSSARY - PERSONS", and "GLOSSARY - CONCEPTS, EVENTS, INSTITUTIONS". The main content area is titled "INTERVIEWS: [1/12]" and lists four entries:

- Ebinger, Endre** (1932), Teacher, physicist. 1956: member of the National Council, Town of Mosonmagyaróvár, sentenced to 6 years' imprisonment. [Public] >199< 1989. 140 pages and attachment, Adrienne Molnár.
- Farkas, Vladimír** (1925-2002), 1939-1945: lived in the Soviet Union. 1946-1955: lieutenant-colonel, State Security Office. 1956-1960: imprisoned. [Closed] >200< 1988. 300 pages and attachment, Gyula Kozák.
- Gyenes, Judith** (1932), Librarian. Widow of Pál Maléter. Dismissed from her job and spent several years as an unskilled worker. [Public] >694< 1998. 100 pages, Adrienne Molnár.
- Horváth, Csaba** (1932), Worker, economist, translator. 1956: Chairman, Workers' Council of SZIM Machine Works in Esztergom, 1958: sentenced to 10 years' imprisonment. [Available for research] >236< 1990. 300 pages and attachment, Adrienne Molnár.

contemporary history. The other elements of the database include the photo archive (some 15,000 digitalized photographs with technical and content descriptions); the historical chronology; the biography archive; the trial documents archive; the video archive; and the bibliography. Each of these database elements is closely associated with the interview subjects. There may be a photograph of the person interviewed, a biography of the subject, photographs of events, or a chronological description.

#### **'Private history' content: '56 and the Kádár period**

Based on these interconnecting database 'elements', we produced a content service entitled 'Private History. 1956 and the Kádár Period'. The content basis for this is a partial set of the interviews in the OHA, the basic data in these interviews, and details and sound documents edited from the full interviews. In these interviews, the subjects speak from the viewpoint of people once condemned for their actions during the Kádár period. The Internet content service also demonstrates the connections between the interviews and persons, perhaps linking an edited biography from the biography archive with photographs from the photo archive. Based on the database, we are able to produce a book-like block of information via the Internet, whose 'technical organizer' is a subset selected for a specific attribute of the elements of the interview set. Within this system and based on such attributes, successive new 'books' will be possible.

#### **Main features of the framework system:**

1. The system allows 'almost complete' digital archiving of the documents, based on Oracle database handling software. The internal database maintenance program allows for the loading of the full texts of documents (which will also be archived), digitized files of related audio-visual documents (sound or video recordings), details of the full text edited according to various criteria, and descriptive data on the documents' content and technical characteristics.
2. It provides an immediate framework system for Internet publishing, fine tuning of a grading system for user accessibility (fully public, or full access only to researchers or staff of the data owners) and secure handling of 'sensitive' non-public data.
3. In developing the data structure of the documents, attention was paid to offers to do with qualitative data archiving from DDI, Dublin Core and TEI, and it is suitable for preparing XML outputs compatible with further development of the system.

#### **Problems and solutions during development:**

1. The life interviews were conducted in Hungar-

ian, which confines them to a relatively small audience. However, the basic data in the interviews and the biographies of the subjects have been translated into English.. We have also translated some of the interviews themselves. Of course it would be desirable in the long term for such tasks to be performed by an automatic translating program, but such devices are not yet sufficiently developed for long texts in the Hungarian language. Native English-speaking translators are being used for the time being and the translations are loaded into the appropriate English fields. (Developed translating programmes would confine the native-speaker translators to checking and editing the texts linguistically.)

2. The life interviews contain references to many events. These can sometimes be linked with actual chronological events, but in each case, what appears is a subjective interpretation. The connection with the actual event has to be signalled, but the data discrepancies also need to be recorded.

3. The many connections between the edited interview texts and other elements of the database (photos, biographies, events) should ideally be found automatically by the program among the requisite objects in the databases, allowing for further exploration through these references. Some of the photographs have been linked singly to the appropriate references, but these references are also embedded in the database. There is a complex problem with footnotes to interpret and explain the text. In some cases, a section of the text has to be explained, elucidated or possibly interpreted. Such footnotes are individual and of no use elsewhere in the text. Other sections of the text call for footnotes that expand on a concept, institution, event or person. These can be applied elsewhere in the text as well and often exist already as data records elsewhere in the database (as entries in lists of concepts or biographies or in a chronology etc.) In the latter case, a cross-reference is needed to the appropriate record. This facility has not been created under this project, partly for technical reasons, and partly due to incompleteness in the content. However, the problem has been addressed with data storage between meta-references in the individual footnote texts. It may even be possible to pick these passages out automatically from their texts, during later development (for instance, with a separate data table).

4. Some parts of the database (OHA interviews, trial documents) also contain confidential information. It was very important for the internal data maintenance module to operate according to a well-defined, graded user-entitlement system, so that such information cannot be accessed from the Internet in any way.

5. The interview texts and interview subjects provide much data of a statistical nature for qualitative sociologi-

cal researches. We would like to ensure the possibility of making the data available to various analytical programs partly through requisite exchange formats (with coding if need be) and partly directly. For the time being, direct access is ensured for programs capable of linking directly with the database. (We are experimenting with Mineset.)

6. The problem of data archiving. The really desirable solution would be for all objects for archiving to be loaded into the database. However, the databases still contain only 'simplified' versions of the digitized photos and sound documents. Full-sized digitized photo files with authentic colour can approach 100 MB in size, not to mention the size of long sound documents. These original files are kept at present in external data stores (CD-ROM, DVD-ROM, DLT). Photographs are available within the database as 300 x 300 thumbnails and 600 x 600 viewing-size pictures. The sound documents are stored in highly condensed WMA format versions.

7. We intend to install an XML format output for connecting with other databases, for data conversion and for data exchanges. This is not yet ready. We are seeking other institutions to provide similar historical or social scientific content, so that we can develop a modern-history portal service that is as broad as possible or join a similar service ourselves.

### Acknowledgements

The inspiration for the online multimedia database of the 1956 Institute's Oral History Archive (<http://server2001.rev.hu/oha/index.html>) was the Edwards Online project (<http://www.qualidata.essex.ac.uk/Edwardians>). Special thanks are due to Marcia Freed Taylor, director of ECASS, for providing a scholarship in 2003, and to Louise Corti, associate director of UK Data Archive, for arranging, during the scholarship period, for me to examine digital archiving and content service practice of the UK Data Archive and Qualidata. No less important to implementing the project were the contributions of the Oral History Archive staff—the framework system for the digital content service crystallized in its final form after long, intensive and fruitful discussions with them. Of the Institute staff, thanks are due in particular to Adrienne Molnár, head of the Oral History Archive, and to Judit M. Topits, who acted as the daily operative coordinator and organizer for the whole project.

### References

1. Edwardians Online: <http://www.qualidata.essex.ac.uk/Edwardians>.
2. Baker, Emma J., and Louise Corti, 'Edwardians Online'. In: *IASSIST Quarterly*, 26:4 (Winter 2002), 5–7.

3. DDI. <http://www.icpsr.umich.edu/DDI/ORG/index.html>.

4. Kőrösi, Zsuzsanna, and Adrienne Molnár, 'Carrying a secret in my heart...'

5. Children of the victims of the reprisals after the Hungarian Revolution'. In: *1956*. Budapest/New York: Central European University Press, 2002, 200 pp.

<sup>1</sup> This work was done with financial support from the Improving Human Potential (IHP) and Knowledge Base-enhancing Access to Research Infrastructures programmes. A research scholarship was provided in 2001–3 by the European Centre for Analysis in the Social Sciences (ecass).

\* Paper presented at the IASSIST Conference, May 2004, in Madison, WI, USA. Zoltán Lux, Institute for History of the 1956 Hungarian Revolution, H - 1074 Budapest, Dohány u. 74., email: [luxz@helka.iif.hu](mailto:luxz@helka.iif.hu), URL: [www.rev.hu](http://www.rev.hu).