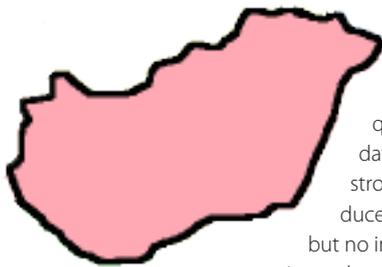# Qualitative Longitudinal Research and Qualitative Resources

The Hungarian Case by Judit Gárdos and Gabriella Ivacs[1]

Abstract

In Hungary, like in many of the former socialist countries, regardless of ongoing research projects based on qualitative data, archives of such research data are sporadic and fragmented. The strong tradition of sociological research produced a vast amount of qualitative social data but no infrastructure was developed to capture, organize and make available these data sets for further analysis. In most cases the secondary analysis or the long-term preservation of these valuable collections is hindered by institutional and disciplinary boundaries as well as a culture of unwillingness to share data in the humanities and social sciences. In Hungary future progress can only be expected if legal barriers are removed or addressed at the policy level; if the underlying infrastructure offers an attractive set of tools for future data collectors; if potential researcher's become aware of the improved access to data; and finally, if preservation best practices, standards, and benchmarking are spread among custodians of the content. It remains an open question which approach is more appropriate: creating a central data repository or setting up a decentralized network of institutions and individuals which can lead to inter-operable platforms to share content and know-how.

Keywords: fragmented archives, legal barriers, data preservation, lack of benchmarking, Hungary

Background

In the past 50 years in Hungary, tens of thousands of interviews have been conducted across widely divergent topics of sociology. This giant scientific source is scattered, idle and is gradually perishing. Even though publications and articles referring to the original data sets do not cease to come out, the raw data is unlikely to be found with ease. To our knowledge, surveys, interviews, and transcriptions collected by researchers were often merged into personal collections of known scholars. Sometimes these files were simply discarded due to lack of space, preservation problems, or, in more fortunate situations, they were donated to libraries, archives or museums but without being described and made available for research. Inadequacy of cataloguing and preservation by all three types of institutions is due to their lack of expertise to describe these materials in depth, e.g., archival descriptions do not go beyond the functional grouping of files. For example, before the Transition period the research projects related to Roma minority issues in Hungary were conducted under various umbrella projects for fear of exposing serious social issues. In the Communist era terms such as Roma problems, unemployment, and social integration were not part of the official research discourse, which was controlled and censored by the Party. Moreover, legitimate research results had to be hidden in the drawers of dusty filing cabinets, and published reports had to avoid sensitive issues as well.

In the past two years, the pilot project called "Voices of the 20th Century - Archive and Research Center" (For further information on VOICES see, www.voicesofthe20century. hu), financed by the Hungarian State Research Fund, has been creating in an inventory of existing resources of qualitative social scientific data, especially interview material gathered in Communist and Post-Communist periods. This comprehensive inventory (now freshly available on the website) is the first step towards setting up an open and public online archive. Our aim is to make the textual and audio-visual heritage of the history of the 20th century accessible to broader audience, both scholars and the

public. Broad public access would enable informed citizens to learn about the 20th century and experience it in a sympathetic way. On the other hand, advanced users of the material – researchers, teachers, media workers and students – could analyse the context in which data was collected, how the results were built into scientific knowledge, and how they could be re-assessed to generate new scientific knowledge. Initially, qualitative audio and audiovisual social scientific collections are being uploaded to the webpage.

It is also to be noted that segregated resources based on disciplinary silos can offer less value to researchers who are interested in broader subjects, periods, and phenomena, rather than a particular social data set from a given period. It is highly recommended to evaluate the possibility of integrating qualitative data archives in Hungary into the existing archival infrastructure, if an adequate one already exists. Archival management, preservation methods, and technology issues can be easily adapted to the special content; in exchange, integrated resources are more attractive to researchers.

For archives and information professionals it is always a challenge to work with legacy data; contextual information attached to collections gathered for various purposes in various ways is as valuable as the records themselves. Preserving provenance information will remain a real problem in the case of scattered qualitative data sets in Hungary as well. The chain of custody in many situations is not clear: nobody knows how certain sets of files migrate to particular members of research teams, who holds the most comprehensive collection, how many copies exist with whom, and so on.

At the same time ownership and intellectual property rights issues also ought to be addressed: research projects financed by state funds are often considered private initiatives, and regulations imposed by the funding organs to share collected data with 'secondary' users are rarely enforced. Secrecy, restrictiveness, fragmentation, isolation, and lack of awareness of cooperation, sharing and preservation continues to characterise the Hungarian research culture, and we can state that in the last 50 years no real progress has been made in this respect. Meanwhile quantitative social sciences and quantitative data archiving are in a somewhat more favourable situation. There is no the doubt that the internationally known institution, TÁRKI Social Research Institute (the only Hungarian member of CESSDA) has had a crucial role in introducing the practice of re-using quantitative data from empirical social research, but the use of their archive is not widespread either.

The Hungarian Data Protection Law seems to be another obstacle to interdisciplinary data sharing: the current version of the legislation does not provide straightforward guidelines on how to make digitized resources containing private data available online, how to define private data in the technology enhanced environment, or how to anonymize digital content in a less difficult way. The Law is deeply rooted in the outdated structure of traditional paper archives. A model of distributed archiving, sharing a common infrastructure and with the participation of several players from organisations to individuals is not approved by the Legislation.
Rum harupic imusdae perchil iquaeped et landaepro delest molorat aquiatiam remquas ventiberiti vollupt atiassimpere voloruntus pero-ratem invelecta cor samendi a que quatibe rsperrovit

The VOICES project
The project is overseen by the Institute of Sociology of the Hungarian Academy of Science, in cooperation with the Open Society Archives at Central European University and the Hungarian National Audiovisual Archive. The VOICES team has developed a two year project plan ending in spring 2011 to examine through a pilot how different types of interview material can be re-used for further research, with their various formats enhanced by technology solutions such as digitization, OCR, voice recognition techniques and so on. Having completed the pilot, a sample collection containing mixed media are being published online along with the contextual information about the methodology employed during the data collection process. Prospective students, researchers of science studies, cultural studies or social studies will be involved in the early project stages to explore the content and incorporate the data into their early works: theses, publications, and other projects. Other expected outcomes of the pilot project are a mechanism to regulate online access according to privacy levels, added features like tagging, annotating, searching and browsing on the web site, and published resources on the methodology of research.

The future archive has three pillars: preservation, research and access. As a result, the VOICES team has direct and continuous contact with new qualitative research projects and collections as well as undergraduate, graduate and post-graduate university programs. The planned website gives an opportunity to co-operate with similar domestic and international initiatives, e.g., other sub-archives, user networks, links, and partner institutions (VOICES is now part of the EQUALAN network of European Qualitative Archives). Our research team, being well-integrated in international research and archival networks, will be able to inform the wider international scientific public about the work of the Voices of the 20th Century Archive.

A plan about data protection and confidentiality for our new archive has been completed (in Hungarian), in accordance with the Hungarian Ombudsman for the Protection of Personal Data. Such a plan proved to be crucial in gaining state funds, since there has been considerable concern about data protection in the reviewers' reports. Data protection is a sensitive issue in many post-socialist countries, Hungary among them. In particular, the role of key personalities in the former socialist secret service has been a huge topic in the Hungarian public sphere. Some historians–using state archive material–reveal from time to time the real names of secret service agents, some of them prominent individuals in contemporary politics or arts. This is made possible by a law text (§32 of the data protection law) stating that if it is necessary for the description of the scientific results concerning events of a historical period, personal data can be published without the consent of the person. In some cases, there has been considerable public discussion whether the identification of a specific person was really needed, and there have been lawsuits against historians for publishing personal data.

Lately, there have been lively discussions among quantitative social scientists about to re-use and merge statistical databases generated in the administrative branch (government organizations, municipalities, censuses, etc.). The question of anonymity is crucial here, since two anonymous databases, if combined, can possibly lead to a new one in which identities can be traced. There is the desire for countrywide guidelines concerning the merging of mainly state-owned databases, and short proposals for a new law have been issued in May 2010 by the Fényes Elek Research Center for Statistics and Econometrics of the Hungarian Sociological Association.

VOICES in the context of archiving in Hungary
This is the context in which the VOICES Archive has to be established. We have been able to investigate the attitudes of some other archiving

institutions and we are becoming aware of the fears of individual researchers in Hungary, which are yet to be fully understood; we have just begun our work with them. The institutions are very protective about their material. There seem to be several reasons for this attitude. For the big state archives such as the State Radio, easy (online) access even to the catalogue is not an important principle. They seem to be content with their huge holdings and are not very eager to cooperate substantially either with other institutions or with individual users. The Audiovisual Archives of the National Széchényi Library cite reasons to explain their data protection policies: there is no existing public catalogue of the interviewees—although such a restriction is not required by the current laws. The researcher can find out whether a certain person has been interviewed, but he cannot ask for a complete list of all interviewed persons. We have only some numbers available: The Collection of Historical Interviews in this Archive has collected and produced approximately 1500 interviews with the principal figures of Hungarian history, public life and culture since 1985.

In the case of research institutions, we encountered a somewhat similar attitude, maybe typical of countries with weak qualitative archiving and sharing traditions. Research institutions seem to guard their findings and are not particularly supportive of "alien" researchers. There are some peculiar means used to restrict information about their material: in the case of a research institute with extensive oral history interview materials, the catalogue describing which interviews are also available in audio format is not made public. But if a researcher asks for a particular interview, he or she will be able to get this information. There seems no rational reason for such conduct. These different attitudes of researchers and institutions would make an interesting research topic and would reveal much about the post-socialist research culture.

Archival holdings in Hungary

Most archives collecting scientific interview material focus on topics related to the humanities, especially World War II and the socialist era. The first topic is represented in Hungary by Centropa, an interactive database of Jewish memory. The project combines old family pictures (a database of 25,000 digitized images) with their accompanying stories (abridged and summarized interviews presented as texts on the homepage and parts read by an actor). Centropa has interviewed more than 1,350 elderly Jews living in Central and Eastern Europe, the former Soviet Union, and the Sephardic communities of Greece, Turkey and the Balkans. It has a separate Hungarian sample and institution. The second topic is covered by the Oral History Archive (OHA) of the Institute for the History of the 1956 Hungarian Revolution, a collection of oral history interviews focusing on the Revolution and the era of János Kádár. The Open Society Archive (OSA) also holds data on the socialist era and archives research resources for the study of Communism and the Cold War, particularly in Central and Eastern Europe, as well as issues of human rights. The OSA has launched a Digital Archive, with the aim of broadening access to primary sources by overcoming technical, legal, geographic, and socio-cultural barriers. They developed a strategy which includes large-scale digitization, multilingual description, and the implementation of open-source solutions and open standards. They also seek to meet current international benchmarks by becoming a trusted digital repository.

To our knowledge, there are no qualitative longitudinal data collections in Hungary. All the above mentioned collections contain material addressing life courses and oral histories. There are some follow-up studies, but only on the level of the individual researchers. One of the aims of our new Archive VOICES will be to investigate who has done qualitative longitudinal projects, as they are not sufficiently documented even at the level of scientific publications.

The only Hungarian institution that is a member of CESSDA is TÁRKI. The TÁRKI Data Archive is the national data repository for empirical research data in Hungary. It collects, stores and publishes a great number of surveys for the social science and business communities nationally and internationally, focusing on quantitative data. The aim of VOICES is to create an internationally acclaimed research database for qualitative social scientific resources in Hungary as well, since –unfortunately– there are currently many gaps in qualitative archiving in Hungary.

The main priorities of VOICES for the next three years are to map existing qualitative resources and to develop a data workflow of creation, capture, deposit, archiving and long-term curation based on international standards. We are engaged in exploring cost-effective technological solutions for managing digital data, in building a network of partners (such as potential data donors, experts and users) and in shaping the Hungarian research culture of not sharing data.

There are some major problems to be solved during our work, such as gaps in Hungarian legislation and outdated legislation, IPR issues, the handling of personal data, the lack of a community approach in the sharing of scientific data, and of course the shortage of long-term funding and the need for sustainability.

Existing organisations (IASSIST, CESSDA) might be of assistance in many ways, among them with their know-how on inventorying, collecting data, and depositing content, their expertise in data archiving systems, technology solutions and in the methodology for developing the policy framework for managing and redistributing data sets.

To secure their future, qualitative archives all over Europe depend on constant and reliable financial help. Being embedded in a community of archives brings new opportunities for mutual assistance in issues of both science and funding.

Further reading
For a bibliography about oral history literature in Hungarian language see: www.replika.hu/system/files/archivum/58-07.pdf

One of the rare analyses of Hungarian oral history: Lénárt, A. (2007). Történetgyűjtés. Az Oral history tudományos műhelyei Magyarországon 1945 után. [Collection of History. Oral History Labs in Hungary after 1945] Aetas 2. pp5–30.

Székely, I. (2009.) 'Positive Discrimination and Data Protection: A Typology of Solutions and the Use of Modern Information Technologies'. In: Szabó, M. (ed). Privacy Protection and Minority Rights. Eötvös Károly Public Institute, Budapest. 27-62.

Note
1. Judit Gárdos Hungarian Academy of Sciences, Institute of Sociology, "Voices of the 20th
Century - Archive and Research Center", The work of Judit Gardos has been partly enabled by fund no. 77566 of the
Hungarian Scientific Research Fund (OTKA). gardos.judit@socio.mta.hu

Gabriella Ivacs Open Society Archives, Central European University, ivacsg@ceu.hu