

Report on Destroyed Communities Project — January 2008

The purpose of this multi-year project is to use Yad Vashem's Central Database of Names in conjunction with other resources and innovative computer technology to develop a family tree for each individual cited in the YV database. On a community-by-community basis, the family trees will be merged with other relevant data to recreate in a virtual environment the network of kinships that characterized these communities on the eve of the Shoah. Work will proceed on two tracks. Track one will enlist the active participation of nonprofessional genealogists researching specific towns. The second track, representing the major effort, will employ experts to develop software needed to analyze the Yad Vashem database as a whole.

In 2006/7, a manual exploratory study was undertaken of one community in an effort to identify unanticipated issues and problems. A technical committee, headed by Jean-Pierre Stroweis was assembled and met in New York in August to discuss an overall approach to development of the needed new software. Stroweis also met with personnel at Yad Vashem to identify areas of mutual interest and cooperation. Yad Vashem has provided its complete data on selected communities. It also undertook to provide its list of alternate community names. Prof. Ruvin Ferber of the University of Riga, has expressed interest in using his complete data on Latvian Jewry for a joint effort. Professor Daniel Wagner has undertaken to develop a segment of the needed larger software, using vital statistics and cemetery data from Ostrov Maz., Poland.

In June 2007, the Institute awarded Destroyed Communities project received a grant of \$10,000. Prof. Wagner also received a grant of \$10,000 for his project

On July 28 and 29, 2007, a subset of the technical committee held an intensive two-day meeting in Newark, New Jersey, to develop a feasibility study for the software needed for track two. Possible funding strategies also was addressed briefly. Participants included Alexander Beider (Paris), the world expert on Jewish names, Dr. Stephen Morse (San Francisco) developer of the Pentium chip, Jean-Pierre Stroweis, team leader for software development, Israel Aircraft Industries, Logan Kleinwaks, an extremely talented math and computer expert, Gary Mokotoff and Sallyann Sack.

The committee concluded that it likely would be possible to develop the software. It split into two teams of two each to work on the two separate, necessary aspects. Alexander Beider and Steven Morse are working on the development of algorithms for phonetic matching of names; Jean-Pierre Stroweis and Logan Kleinwaks are tackling the development of algorithms that will permit identification of relationships between individuals. Demonstrable progress has been made by both subcommittees.

The committee also recognized that additional resource tools will need to be developed, specifically dictionaries of Jewish names similar to the ones that have been written by Beider for Russia, including one for Romanian names, and another for Hungarian names. The Romanian study currently is being prepared with supervision from Beider. The Project Leader, Dr. Sack, is talking with a potential team for the Hungarian book. Dr. Sack also has been investigating

possible sources of funding necessary to develop the software fully.

The committee will meet again the first weekend in April in suburban Washington, D.C., to demonstrate and appraise the algorithms and models that have been developed thus far and to map the team's next steps, both technically and financially. The balance of the original \$10,000 grant will be used to pay the cost of this meeting.

Sallyann Sack,
Washington, DC.