Digital Library Program Will Access Russian Periodical Index

The Indiana University Digital Library Program received one of eight grants awarded nationally in 1999 under the U.S. Department of Education’s Technological Innovation and Cooperation for Foreign Information Access Program of Title VI. An aim of this grant program is to support projects that can develop techniques to access or preserve foreign information resources in print or electronic form, particularly resources written in non-Roman language alphabets, for wider dissemination.

IU was awarded $240,000 for the “Russian Periodical Index Digital Project.” This 3-year project will digitize a 20-year interval (1956-1975) of the major Russian index of journals and periodicals, Letopis’ Zurnal’nykh Statei (Chronicle of Journal Articles), and make the index available to scholars via the World Wide Web. This government-published index is “a comprehensive, national bibliography of the Soviet Union... that, in its paper form, is hard to obtain and poorly organized,” says Kristine Brancolini, acting director of the Digital Library Program, who wrote the winning proposal along with area librarians Murlin Croucher and Perry Willett. The Digital Library Program is a university-wide partnership of the IU Libraries, University Information Technology Services, and the School of Library and Information Science. IU is a considered a leader in mounting large digital collections on the World Wide Web.

The Russian index is published on highly acidic paper, which becomes brittle and yellow and deteriorates over time. One of only 15 to 20 universities in the United States to have back issues or current subscriptions to this serial publication, “Indiana University was perfectly suited to do such a project,” says Brancolini, citing the academic strengths of the Slavics Department and the Russian and East European Institute, the breadth of the library’s impressive Slavics collection, and the technical expertise of the Digital Library Program staff.

The index is comparable to the Readers’ Guide to Periodical Literature, the Humanities Index, and the Social Sciences Index all combined into one and contains information on articles appearing in more than 1,700 popular and scholarly journals. “There are no subject or title indexes to this massive publication,” says Croucher, Slavic Studies bibliographer. “Anyone wishing to do Soviet research has the daunting task of endless searching to find periodical literature. Digitizing this massive publication will make possible author, title, and keyword searches, instantaneously covering a 20-year run of Soviet journals.”

The period 1956-1975 was chosen not only for its historical importance but also because there already exists a European reprint of the index for the years through 1955. Given the budgetary and time considerations of the IU project, digitization could only be carried out through 1975. That still comprises an estimated 250,000 pages to be scanned by optical character recognition software and rendered searchable.

There are many challenging aspects to such a project, according to Andrew Spencer, a visiting assistant librarian who has been hired as project manager. Spencer, who holds a master’s in library science from IU and a master’s in Soviet and East European Studies, says that the Cyrillic text and the inflectional structure of Russian pose their own special linguistic problems in keyword searching. More technical challenges include assessing the effect of yellowed brittle paper on the accuracy rate of character recognition software and its consequences for manual editing and proofreading; deciding on which markup language works best with the Cyrillic alphabet; finding the optimal resolution of scanned images; building the structure of the database itself; determining the limitations of automating the markup process for such a massive and complex amount of data; and choosing a suitable search engine to make the database keyword-searchable.

In April, Spencer will present a paper at a conference, Managing the Digital Future of Libraries, to be held in Moscow at the Russian State Library where he will discuss a number of the preservation and technical issues associated with the Russian digital project. For further information, contact Andrew Spencer by e-mail at gaspence@indiana.edu or by telephone at 812-855-1261.

—RMN