

Editor's Notes

Welcome to the second issue of the IASSIST Quarterly, vol. 30.

Winter came late in Denmark, but suddenly the situation was normal: snow came and traffic stopped. And now 14 days later spring is here. I sat outside in the sun reading articles. Still it's more difficult to write articles outside; we are waiting for improvements to the computer screens. However, some people have stayed inside to write articles for the IASSIST Quarterly. They are presented below.

The first article, "Microdata Information System - MISSY," is written by Andrea Janssen and Jeanette Bohr from Centre for Survey Research and Methodology, ZUMA, at Mannheim in Germany. The MISSY was presented by Andrea Janssen, Jeanette Bohr, and Joachim Wackerow at the session on "Effective Strategies for Metadata Management" at the IASSIST conference in May 2006 in Ann Arbor. The data in the system are from the German microcensuses for 1995 and 1997, which contain a sample of one percent of all German households. The microcensus has been carried out since 1957, and parts of the microcensus are available for research. The researchers need extensive metadata on both the study and variable level, e.g., the microcensus uses complicated classifications of professions, sectors, and household arrangements. The system is based on the standard from the Data Documentation Initiative and documentation includes general information such as questionnaire, codebooks, interviewer guides, frequencies and also some tips and recommendations on use of the data. MISSY is a German system, in German language, and for researchers in Germany. However, all are free to gain from the experiences presented in the article.

The second article was also presented at the IASSIST 2006 conference, at the session "Innovations in Data Dissemination." The title of the article is "User-Centered Design and Innovation in the Sociometrics Social Science Electronic Data Library (SSEDL)." The authors, Josefina J. Card, Tamara Kuhn, and Thomas Wells, are all at Sociometrics Corporation. The article describes the Sociometrics Data Archives SSEDL as being a rich source of data for those in the public health, medical, nursing, social work, and social science professions. In the "product package," datasets come with several data input files for SAS and SPSS. Purchasers can acquire data on CD-ROM or download from the web site, and downloading has had a significant rise in usage. In order to help users identify relevant datasets, Sociometrics has launched a topic-based drill down system showing "areas of richness," which helps users identify and reach datasets and variables and the appropriate documentation.

The last article is "Overview of a Proposed Standard for the Scholarly Citation of Quantitative Data," by Micah Altman and Gary King from Harvard University. This is an extended abstract summarizing a proposed standard for citation. This was presented at the IASSIST 2006 conference at the session "New Standards in Statistics and Data Citations." The authors mention that, at a minimum, citations should include author(s), date of publication of the data set, and the data set title. These fields have been discussed before during the close to 50 years of documentation of data sets, and the fields are not so unambiguous now not to be discussed further. For instance, a great many people can be called "authors" in the production of a data set, and the same data set can have several other relevant dates attached besides "publishing date." The authors recommend using the DDI elements, but their main purpose is to propose some novel fields that are directly linked to the use of modern technology. First of all, a Unique Global Identifier. The authors mention a naming resolution service and that brings to mind the technology of the Internet with name servers for looking up the correct IP address; however in this context, more than one copy of the dataset can exist at different physical locations. Secondly, technology is applied by adding a Universal Numeric Fingerprint. This should guarantee that the data set has not been changed even though the data set might exist in different software. This should probably apply to the documentation as well.

The IASSIST is always open at its website, <http://iassistdata.org>, where you can look at conference information and visit the IASSIST blog (IASSIST Communiqué, <http://iassistblog.org>).

Articles for the IASSIST Quarterly are most welcome. Articles can be papers from IASSIST conferences, from other conferences, from local presentations, discussion input, etc. Contact the editor via e-mail: kbr@sam.sdu.dk.

Karsten Boye Rasmussen, March 2007