

## Project Summary: A Remote Access Analysis System for the Survey of Living Conditions in the Arctic

A major challenge facing Arctic social scientists is making survey data sets available to the research community and indigenous regional organizations while maintaining the confidentiality of survey respondents. The reliability of survey results is directly related to sample size. This means that it takes the same size sample to describe Inuit in the Arctic as it does to describe all people living in the Arctic. A sample of approximately 625 people, for example, is required to describe a characteristic of Inuit adults with a sampling error of plus or minus four percentage points. In Alaska, just over 11,000 Inuit 16 years and older live in the three Inupiat Settlement Regions (North Slope, Northwest Arctic, Bering Straits). Inuit adults in these regions have one chance in twenty of being selected in a sample of 625. Given such high probabilities of selection, one can easily imagine that it is possible to identify (in a fictitious example) the one of twenty individuals interviewed in Kivalina who reported their occupation as a health aide. Identification of respondents would not only violate the conditions of informed consent; it would also put respondents at risk as the interview included highly sensitive questions about such topics as drug use, domestic violence, and suicide.

The international Survey of Living Conditions in the Arctic project team is proposing to collaborate with the University of Michigan Inter-university Consortium for Political and Social Research (ICPSR) and the Berkeley Survey Documentation and Analysis (SDA) program to make available for further analysis the results of a representative sample of indigenous adults living in Inuit settlement regions in Alaska, Canada, Greenland, and Russia. The data file consists of over 7,000 interviews. Building on the concepts of **remote access analysis**, the team will provide researchers and policy analysts with the capability of applying a full array of statistical analysis techniques to the dataset while ensuring that the individual records remain unseen. The remote access analysis system will also be programmed so that researchers cannot inadvertently request analyses that could reveal the identity of individual respondents.

The original intent of the Survey of Living Conditions in the Arctic project team was to make the full data set available on a CD after removing all individual identifiers and collapsing response categories to remove the risk of identifying individual respondents. The analysis experience of the project team has convinced the team that the CD dataset approach would require such extensive grouping of response categories as to severely limit the range of research questions that can be addressed.

An alternative approach to that of providing a CD is to make the dataset available within one or more secure analysis laboratories. The SLICA team is currently conducting its international analysis within such a laboratory system in Canada. This approach, however, is costly to researchers in terms of both travel expenses and time. The secure laboratory approach is also a more labor intensive system to maintain. An additional benefit of the remote access analysis system is that it will be available to researchers and analysts through a secure internet connection. Thus widely dispersed Arctic indigenous organizations and researchers will be able to conduct individual analyses and more easily collaborate without incurring travel expenses.

**Intellectual Merit:** Data produced by the Survey of Living Conditions in the Arctic represents a unique resource for scientific inquiry. It includes comparative data for regions in Alaska, Canada, Greenland, and Russia. The dataset consists of individual records of responses to a comprehensive array of living conditions, examples of which include health, hunting and fishing, employment, and civic participation. The data can be used by researchers and analysts to address both theoretical and policy questions. The intellectual merit of this proposal is to foster further research by making the dataset available to the research and policy analysis communities.

**Broader Impact:** A remote access analysis system can both limit the risks of disclosure of the identity of respondents and foster use of the international dataset. Its broader impact will be to both provide research opportunities for graduate students and to allow indigenous organizations and other policy organizations to inform themselves about living conditions among Arctic peoples. The project's broader impact will also be to address the more general question of how to archive and make available for access survey datasets involving samples of small populations and collection of highly sensitive information.