

Dated March 27, 2007

**Research Data Technology Centre  
Advisory Council**

**Minutes from Meeting  
23 March 2007**

**Council:**

Roderic Beaujot, Sociology (Academic Director)  
David Bellhouse, Statistical and Actuarial Sciences (regrets)  
Joyce Garnett, Chief Librarian (regrets)  
Iris Gutmanis, Evaluation and Research, Parkwood Hospital  
Tedd Hewitt, Research V-P (Alison Ariss as replacement)  
Julie McMullin, Associate Dean, Faculty of Social Science  
Reza Nakhaie, Sociology, (Daniel Edelstein as possible replacement), University of Windsor  
(regrets)  
Chris Robinson, Economics  
Bridget Ryan, Epidemiology and Biostatistics (student representative)  
Moirra Stewart, Family Medicine (regrets)  
Stewart Thornhill, Ivey (Terry Wong as replacement)

**Staff:**

Cynthia Cook, Statistics Canada (Senior Analyst, Regional Supervisor)  
Bradley Corbett, Statistics Canada (Analyst)  
Vince Gray, Social Science Network and Data Services (Liaison Officer) (Doug Link as  
replacement)  
Zenaida Ravanera, Information Officer

**1. Minutes of 19 October 2006**

The minutes have been approved and posted.

Business arising:

Welcome to Bridget Ryan as Student Representative on Council.

University of Windsor is now an official Partner of the Research Data Technology Centre, and listed as a Partner in the national RDC system.

**2. Budget**

The budget page for 2006-07 and 2007-08 was reviewed.

The exact amount remains unknown for the 2007-08 revenue from the SSHRC/CIHR Operating Grant to Research Data Centres National Coordinating Committee. With an estimate of \$63,000 from this source, and depending on the revenue from University of Windsor, the amount from the Faculty of Social Science may need to be raised to \$28,000 from the current \$20,000.

The expenditure lines use the categories of the RDCNCC SSHRC/CIHR grant.

### **3. Activities of the Research Data Centres National Coordinating Committee**

#### **3.1 CFI Leading edge grant**

The NCC has obtained a CFI Leading Edge Grant of \$4 million on a \$10 million budget. This is administered by Raymond Currie at University of Manitoba. This grant will cover (1) a high speed Virtual Private Network across the RDCs (still to be approved), (2) development of meta data on data sets, (3) equipment renewal across the RDCs later in the grant.

#### **3.2 Next NCC meeting at Western**

The 26-27 April meeting of the NCC will occur at Western, using the facilities at Windermere. There will be a small open house permitting the NCC to visit our RDC, to which members of Council will be invited, around 5:00 pm on 26 April.

#### **3.3 Demonstration of IDLS within RDC**

Doug Link presented the outlines of a demonstration that will be presented to the NCC, regarding using the Internet Data Library System (IDLS) within the RDCs. This had been on our CFI application for the Western Research Data Technology Centre, and it now becomes more attractive as a virtual private network across the RDCs is being considered. Since the 1970s, the Social Science Network and Data Services (SSNDS) has developed and supported data and metadata search and retrieval systems which now serves 37 of the approximately 70 DLI member institutions. A letter of understanding is in draft with the Council of Rectors and Principals of Quebec Universities (CREPUQ) for an IDLD-Sherlock partnership project. A project is underway to implement the IDLS as a separate and dedicated server within the secure environment of the UWO RDC. On 26 April, this will be demonstrated as a potential tool over an RDC Virtual Private Network (VPN). The IDLS systems and development team, headed by Vince Gray (data expert) and Les Flodrowski (systems expert) is examining the design and work required for a phased in approach to DDI compliance. The similarity of data delivery in DLI and RDC would permit a quicker move to the analysis stage once a researcher is within the RDC.

#### **3.4 National Knowledge Transfer Officer**

Heather Juby, located in Montreal, is employed as National Knowledge Transfer Officer. She is coordinating with the NCC, the various RDCs and Statistics Canada, to maximize the knowledge transfer associated with RDC products. As “deemed employees” of Statistics Canada, researchers with access to RDC data can have announcements of their products in The Daily of Statistics Canada. This requires much advance notice and preparation, which should start by consulting with Zenaida Ravanera, the UWO RDC Information Officer.

#### **3.5 National Conferences**

The National Conferences will be held in Halifax (12-13 October 2007, “Life course transitions of children and youth” see web site for links) and Ottawa (2008). We have proposed Western for 2009 but we may be pre-empted since that would make two in a row in Ontario.

### **4. Report to CFI, NCC, Statistics Canada, SSHRC and CIHR**

The staff have organized themselves in terms of responsibility for the various items of a report that will be used for the mid-term review at SSHRC/CIHR. This replaces the earlier format used in reporting to Statistics Canada.

In the March 2007 report on the RDTTC to CFI, we indicated 65 people using the RDC in 2006: 47 researchers, 1 post-doc, 7 PhD, 6 Masters, 4 staff. We indicated that one person had profited from the research environment at the RDC as he went to a post-doc at the University of Toronto Centre for International Health. We indicated 8 peer reviewed publications for 2006 or forthcoming. We need more information on how the RDC is a source of attraction when faculty and students come to Western, and more information on how the RDC experience profits the careers of students and faculty who leave Western.

## **5. Projects, researchers, products (Zenaida Ravanera)**

A renewed web site is ready to be launched.

We needed to obtain permission to list the projects along with principal investigators and the data sets. As of 21 March 2007, 40 projects gave permission, 4 projects denied the permission to have the project listed, and 10 have not responded. When we have permission, the project is listed on the web site.

A list of 40 products from 2001 to 2007 was distributed. Products that are already in the public domain will be listed on the web site without seeking permission, in the following categories: book chapter, conference paper, journal article, discussion paper series, MA thesis, PhD thesis. The category of “working paper” will be used for products that are not to be listed on the web site, such as papers in submission to journals. It is important that Ravanera have copies of these papers in submission so that she can plan, with the author and possibly the publisher, for knowledge transfer activities associated with the release.

We have built a data base (using the InMagic DB Text works) for the RDC to keep track of projects and products of the investigators.

The first issue of a Newsletter is ready to post on the web. A copy was sent to Western’s Media Relations for possible further distribution.

All are encouraged to send products associated with RDC data, including those that were undertaken at other settings before Western’s RDC opened.

## **6. Presentations, workshops, (Bradley Corbett)**

### **6.1 Presentations**

Russell Wilkins of Statistics Canada, presented on 23 January on Strengths and Limitations of Small-Area Data in Survey Research and Administrative Data: A User’s Perspective.

A presentation by Myles Corak of Statistics Canada on family data had to be cancelled. We will attempt to re-schedule.

### **6.2 Workshop**

A workshop on “HLM and STATA” was delivered on 26 February 2007, with Brad Corbett and David Walters presenting. There were 14 persons in attendance, including 8 who had received bursaries from the RDTTC to attend the workshop. The workshop was well appreciated and the students in particular were appreciative of the bursaries that had allowed

them to attend.

### **6.3 Course on data and analysis**

Brad is talking with various people about the potential niche for a course offered to people of various departments on the use of large data sets. It could either take a simpler version (use of guides and survey documentation, sample design, bootstrapping, managing and linking data, rounding and minimal cell counts) or include more complex questions (different regression models, multi-level models, survival analysis).

Julie McMullin will determine with Graduate Studies how such a course could be listed.

Until a course is structured, some of these things can be done in workshops. There may also be interest in workshops that would give exposure to several major software packages. The general objective is to encourage more people to use the RDC.

### **6.4 Multi-Level Modelling (MM) Research Unit**

There has been discussion of establishing a research unit that would be a centre of expertise on community effects on health and social outcomes. Presentations will be made at Utrecht University in Amsterdam (April 15 – April 17 2007). A first activity would be to bring back some things learned at this international occasion. Another activity would be to apply for funding either as a research development initiative or to provide specialized training in a summer school setting. In the long term, this unit could be a basis for links with various people interested in multi-level approaches and who would want to add community-level variables in their analyses. The unit could also provide statistical consultancy, based on people who would offer their expertise.

### **6.5 Statistical consultant**

This item is on hold. From the point of view of SSNDS, there would be advantages to handle some of the statistical consultancy that is now done on a contract basis.

### **6.6 Exit interview** to determine training and assistance needs (no report)

## **7. Outreach on campus**

### **7.1 E-lists:**

Vince Gray has developed several e-lists:

uwordc is for Council

uwordcinfo is for operations, people join the list through the web site

uwordcpub is for “agents” who would forward information as they determine appropriate, this is especially graduate secretaries.

uwordcevents is for publicizing events, it includes the three above lists plus others who have had contacts with the RDC.

### **7.2 Other**

We are interested in suggestions that would help improve the profile of the RDC, besides the following:

Fall presentation at SSNDS:

Faculty outreach: new faculty orientation, faculty development workshops  
Announcements to graduate methods courses  
In particular, Brad is able to make shorter or longer presentations to interested groups.

#### **8. Western, Household Panel Survey and Simulated Data (Chris Robinson)**

The people from the CIBC Human Capital and Productivity Centre in the Department of Economics have been selected to develop the human capital module for the new Canadian Household Panel Survey (CHPS) that is designed to succeed the Survey of Labour and Income Dynamics (SLID) as the major Canadian household panel. The RDC at Western will be able to house the pilot data from this survey.

John Abowd's group at Cornell University run a Simulated Data Centre which provides expertise in developing simulated data sets that can be used in cases where there are confidentiality concerns with the underlying data set. Cornell is also the North American base of the Cross National Equivalent File (CNEF) which is designed to provide equivalent data from the major international household panels for use in cross-country studies. (A description of CNEF can be found at:

<http://www.human.cornell.edu/che/PAM/Research/Centers-Programs/German-Panel/cnef.cfm>)

The current Canadian component of CNEF is SLID, but this component is rarely used due to special restrictions on the availability of SLID. CHPS will replace SLID as the Canadian component of CNEF and Statistics Canada are putting a high priority on encouraging greater use of the Canadian data. This will take two forms. First, Statistics Canada is evaluating the usefulness of developing simulated data sets, using the methods of the Cornell group, which could be made widely available without confidentiality concerns. Second, the design of CHPS will include every effort to minimize confidentiality concerns that will allow similar easy access to some form of CHPS as is currently the case for the other international panels in the CNEF. Western's RDC and the Human Capital group are logical partners for Statistics Canada in this process. It would make sense to explore the possibility of a Canadian CNEF base at Western, linked to Cornell to facilitate greater Statistics Canada involvement in CNEF.

#### **9. The "to do" list based on CFI application**

IDLS within the RDC (treated as 3.3 above).

Collaborative joint research link with Simulated Data Centre at Cornell University (treated as 8 above)

Development of event history software (Fernando, no report)

Links with Waterloo, McMaster and Toronto, especially for training and workshops (no report)

Mexico/INEGI (no report)

#### **10. Other business**

nil