
The Canadian Century Research Infrastructure Project and Computing in the Humanities

Chad Gaffield (*gaffield@uottawa.ca*)

Department of History, University of Ottawa

Marc St-Hilaire (*Marc St-Hilaire@ggr.ulaval.ca*)

*Centre interuniversitaire d'études québécoises,
Université Laval*

Claude Bellavance (*Claude_Bellavance@uqtr.ca*)

Centre d'études québécoises (CÉDEQ)

Gordon Darroch (*darroch@yorku.ca*)

Sociology, York University

Peter Baskerville (*pab@uvic.ca*)

History, Univ of Victoria

Introduction

One of the most comprehensive humanities and social science research projects in Canadian history, the *Canadian Century Research Infrastructure* (CCRI) is a five-year, pan-Canadian initiative to develop a set of interrelated databases centered on census records for the 1911-1951 period.

The databases being developed from manuscript census records for the period 1911 to 1951 form the core of a much larger research infrastructure, the objective of which is to construct an evidentiary foundation for research on the transformation of Canadian society from the late 19th century to the later 20th century. To construct this evidentiary foundation, the CCRI will have two major components: primary sources and secondary sources.

Census microdata from the 1911-1951 enumerations form the first and the core of the four primary data sources. Other primary data sources include Statistics Canada documentary sources concerning the enumeration process; newspaper evidence about the enumerations at the time of each enumeration; and House of Commons and Senate debates related to the enumerations. The goal of this component of the CCRI is to provide researchers with the contextual evidence necessary to undertake appropriate analysis of the census microdata.

The secondary data sources are intended to facilitate research on the primary sources, and are equally varied in nature. They range from introductory descriptive statements about the census enumeration process, to highly technical discussions of data-entry and coding issues, and bibliographies of census-research publications.

Integral to the entire project is the construction of a geographic framework for the historical census data, using a *Geographic Information System* (GIS). GIS map layers are being created to enable geographic location, selection, aggregation and analysis of sample data, as well as some mapping of generalized census data. This will allow researchers to ask questions of the database which are much more geographically specific than in the past. Interface tools to make these geographic queries and analysis as user-friendly as possible are also being developed.

The CCRI will be structured in terms of five distinct articulations, each devoted to one of the enumeration years (1911-1951). The CCRI will include cross-census harmonization bridges (or crosswalks) that connect each of the five articulations, to enable comparative analysis. A variety of user guides will be developed to aid researchers. In addition to a general introduction to each census enumeration, there will be user guides for each census variable, as well as a separate guide detailing the coding scheme for that variable. As it is expected that some variables (such as occupation) may be coded according to more than one scheme, each scheme will be discussed in the guides.

The CCRI databases will be made available through Research Data Centres across Canada; versions will also be available through the *Data Liberation Initiative* at Canadian universities. Once completed, the CCRI databases will be joined to other databases that cover the periods from 1871 to 1901 and from 1961 to 2001. The result will be a new foundation for the study of social, economic, cultural, and political change, as the *Canadian Century Research Infrastructure* will include an extraordinary range of data about the twentieth century.

The proposed panel discussion will focus on the following related issues:

- Integrating Words and Numbers in Historical Databases;
- Mapping Time: Using GIS to enhance historical understanding;
- Meta-data, Contextual Data, and User Guides for Historical Evidence: How much is enough?
- Integrating Primary and Secondary Sources in Historical Research Infrastructures.