

HISTORICAL GIS RESEARCH IN CANADA

Edited by Jennifer Bonnell and Marcel Fortin

ISBN 978-1-55238-756-6

THIS BOOK IS AN OPEN ACCESS E-BOOK. It is an electronic version of a book that can be purchased in physical form through any bookseller or on-line retailer, or from our distributors. Please support this open access publication by requesting that your university purchase a print copy of this book, or by purchasing a copy yourself. If you have any questions, please contact us at ucpress@ucalgary.ca

Cover Art: The artwork on the cover of this book is not open access and falls under traditional copyright provisions; it cannot be reproduced in any way without written permission of the artists and their agents. The cover can be displayed as a complete cover image for the purposes of publicizing this work, but the artwork cannot be extracted from the context of the cover of this specific work without breaching the artist's copyright.

COPYRIGHT NOTICE: This open-access work is published under a Creative Commons licence. This means that you are free to copy, distribute, display or perform the work as long as you clearly attribute the work to its authors and publisher, that you do not use this work for any commercial gain in any form, and that you in no way alter, transform, or build on the work outside of its use in normal academic scholarship without our express permission. If you want to reuse or distribute the work, you must inform its new audience of the licence terms of this work. For more information, see details of the Creative Commons licence at: <http://creativecommons.org/licenses/by-nc-nd/3.0/>

UNDER THE CREATIVE COMMONS LICENCE YOU **MAY**:

- read and store this document free of charge;
- distribute it for personal use free of charge;
- print sections of the work for personal use;
- read or perform parts of the work in a context where no financial transactions take place.

UNDER THE CREATIVE COMMONS LICENCE YOU **MAY NOT**:

- gain financially from the work in any way;
- sell the work or seek monies in relation to the distribution of the work;
- use the work in any commercial activity of any kind;
- profit a third party indirectly via use or distribution of the work;
- distribute in or through a commercial body (with the exception of academic usage within educational institutions such as schools and universities);
- reproduce, distribute, or store the cover image outside of its function as a cover of this work;
- alter or build on the work outside of normal academic scholarship.

Exploring Historical Geography Using Census Microdata: The Canadian Century Research Infrastructure (CCRI) Project

Byron Moldofsky

Most readers of this volume will be familiar with the historical censuses of Canada and its predecessor colonies and other regional censuses that were conducted in different parts of the country since 1832. Most will have used the published census volumes for research and reference. As in many Western nations, they are a core source of data for historical research for anyone interested in how, when, and where this nation grew. Less well known, however, are the primary source data for the censuses – the manuscript census schedules themselves, which attempted to record every individual man, woman, and child in the country, usually at ten-year intervals. Although the published volumes contain summary tables derived from these records, the manuscripts may be seen as a vast, scarcely tapped resource for accessing the hidden “history of the anonymous.” Most of these original census schedules – thousands of pages of handwritten enumerators’ log books – have been available only in analog form as images on microfilm in government archives; in recent years, some of the early ones have been scanned and put online for genealogical uses.¹ However, it is still virtually impossible to use these digital pictures of old ledgers for analytical purposes. The “microdata” contained within these manuscript censuses – the information attached to Canadian individuals in

a particular place and time – contain enormous potential for spatial history.

The Canadian Century Research Infrastructure (CCRI) project opens a window into this vast repository of information. A collaborative initiative among fourteen academics at seven Canadian universities, as well as private-sector and governmental agencies, the CCRI has produced a database of a sample of manuscript census data for the Canadian decennial censuses from 1911 through 1951.² The CCRI team elected to focus its efforts on the census enumerations of the first half of the twentieth century – “Canada’s century” – in order to build understanding of the country’s transformation from a “sparsely populated, quasi-colony with a highly dependent economy and settlements spread across a vast territory” to “a highly urbanized and industrialized country on the world stage.”³ These resources are being made available to all scholars for research and provide a core framework that can be built upon in the future.

The project takes as its departure point the concept of research infrastructure: the understanding that the resources underpinning major and collaborative research work can often not be built by individual researchers alone but need public support to provide and enable this public good. The establishment of the Canadian Foundation for Innovation (CFI) in 1997 inaugurated a new approach to academic support in Canada by funding this type of research infrastructure. Initially targeting scientific and engineering fields, the CFI awarded its first major grant in the social sciences and humanities sector to the CCRI project. The project has been the subject of several articles, including a 2007 special issue of the journal *Historical Methods*. Additional background information

and user guides are available on the CCRI gateway website.⁴ This chapter offers an introduction to the project and the resources it offers for historical GIS research in Canada and outlines some of the considerations in working with the data in a GIS environment.

WHAT IS THE CCRI?

The CCRI project consists of three interrelated components: 1) the database of sample microdata; 2) a contextual database of secondary data sources; and 3) a GIS framework that allows researchers to situate the microdata within their geographic context. The sample microdata database incorporates individual records from the manuscript census schedules as recorded by enumerators for the 1911 through 1951 decennial censuses of Canada. The main census population schedules were sampled at rates of 3–5 per cent by dwelling, depending on year (for example, in 1911 a 5 per cent sample was taken, with all individuals in every twentieth dwelling listed).⁵ Every enumeration sub-district was sampled, which in effect provided geographic stratification. These data were entered via a customized data-entry interface and migrated into a relational database for processing. Standardized coding systems based on internationally recognized models were then applied to allow for both easy and sophisticated data manipulation. The main products are five coded data files, one for each census, suitable for working within a variety of statistical software packages. The original database of about eighty hierarchically linked tables, used for data validation, cleaning, and coding, will also

be made available through Statistics Canada for generating more products in the future.

The contextual database includes newspaper and journal writings of the time, documents created by the statistical agencies of the day, and information from the published census volumes. These expressions of public knowledge and opinion about the census reflect the social and cultural context of their times. An important aspect of the project vision was enabling multiple perspectives and views on the censuses from both quantitative and qualitative approaches. Not only the “results” of the census were considered important, but also the “multi-layered political, social, economic and cultural contexts within which enumerations took place.”⁶ As members of the CCRI team have commented, “whereas the microdata derived from the census manuscripts will enable research into the hidden history of the individual lives of Canadians, the contextual data will make possible inquiries about the making and interpreting of that data and the challenges of the great enumerations of contemporary history.”⁷ Thematically indexed and coded, the bilingual contextual database is publicly accessible through a user-friendly web interface.⁸

The GIS component of the CCRI is unique among similar census database projects in its mandate to link the historical census microdata to a geographical framework at a level of detail that was as granular as possible.⁹ This provides major improvements over the conventional mapping of published aggregate census data by standard census divisions (CDs). First of all, the CCRI sample microdata provides access to individual-level data, allowing sub-populations to be extracted for analysis. Researchers can focus, for example, on children, the elderly, women of child-bearing age, or any other

age/gender cohort. Secondly, cross-tabulations of data can be made, such as income levels by ethnic groups. Finally, any of these resulting datasets can then be mapped or analyzed by detailed geographic location. Within the GIS framework, sample data can be aggregated into a variety of geographical “containers,” from standard census subdivisions (CSDs) to customized geographical units, according to the needs of the individual researcher. While issues of sample size, data completeness, and locational accuracy limit to some degree the questions that can be addressed, the CCRI data offer immense potential for historical geographic analysis.

The CCRI builds upon similar work that has been done nationally and internationally to tap historical censuses as an underused source of information about the everyday people who worked, lived, and raised families in different places and times. The project is modelled most directly upon the Minnesota Population Center’s (MPC) Integrated Public Use Microdata Series project (IPUMS), a publicly accessible repository of individual-level sample data from the U.S. census and American Community Survey, from 1850 to the present.¹⁰ The MPC also coordinates the North Atlantic Population Project, which is producing a harmonized database of complete-count censuses from Canada, Great Britain, Iceland, Norway, and the United States between 1865 and 1900.¹¹ The only complete-count census microdata freely available for scholarly research, it is the basis for many international comparative research projects. Several of the principals at MPC provided advice and consultation to the CCRI, and the goal of international comparability was always considered in the decision-making process.

Similar projects in Canada were also influential as models for the CCRI, particularly the 1881 Census project based at the *Département de Démographie*, Université de Montréal, and the Canadian Families project, based at the University of Victoria, which created a similar sample microdata database for the 1901 Census of Canada and from which came several of the CCRI team leaders.¹² Statistics Canada was undoubtedly the most vital external partner on the project, and their Public Use Microdata Files (PUMF) also influenced the data design. Other projects are now underway, which also have commonalities with the CCRI and present the potential for research interaction, including the 1891 census project based at the University of Guelph, and the BALSAC project, jointly operated by four Quebec universities, which links vital statistics records from Quebec.¹³

The CCRI team aspires to make all of the data and documentation as freely available as possible over the web. However, this goal is complicated by the fact that, unlike most of the sister projects referenced above, the CCRI deals with census enumerations whose confidentiality is protected by Statistics Canada's covenant with the Canadian people to respect individual privacy. Historically, an informal tradition developed in Canada whereby individual-level census data were released to the public only a number of decades after their collection. This was formalized in legislation by the *Privacy Act* of 1983, which established the ninety-two-year rule for the release of census enumeration data.¹⁴ In order to compile the CCRI data still protected under the *Privacy Act*, the project established seven high-security data-handling centres across the country,

and all staff were sworn in as deemed Statistics Canada employees.

As a result of these confidentiality rules, and in accordance with different access restrictions, the CCRI employs a range of methods to disseminate its data products:

1. The 1911 sample microdata is freely available publicly as a downloadable microdata file. Full records can be selected and extracted or the full data file downloaded.¹⁵
2. The 1921–51 sample microdata files are currently available only in Statistics Canada's Research Data Centres (RDCs). Like the anonymized microdata that StatsCan creates for recent censuses, these files contain records of individual people, but stripped of personal identification such as names and addresses. Researchers must apply to the RDC program, justifying their need for the data and explaining the nature of their use. The review process typically takes two months. Once approved, researchers may use data only within these high-security RDC centres and must follow disclosure rules that limit what output can be taken away and what may be published.¹⁶
3. The CCRI geographic GIS files (in ESRI shapefile format) and reference maps (in PDF) for 1911–51 are freely available for download from the CCRI geography website.¹⁷ Use of GIS files is subject to Statistics Canada licensing restrictions for 2001

Dissemination Area Cartographic boundary files.¹⁸

4. The contextual database for 1911–51, and its companion user’s guide, are publicly available for download on a bilingual website hosted by the Centre interuniversitaire d’études québécoises (CIÉQ).
5. Two supplementary data products are also publicly available on the main CCRI website: 1) the digitized published tables, which include selected tables of aggregate data from the published census volumes 1911–51 in spreadsheet format; and 2) the CCRI user’s guide, which documents the data design, variables, and coding used in the CCRI microdata files.

CONSIDERATIONS IN WORKING WITH THE CCRI DATA

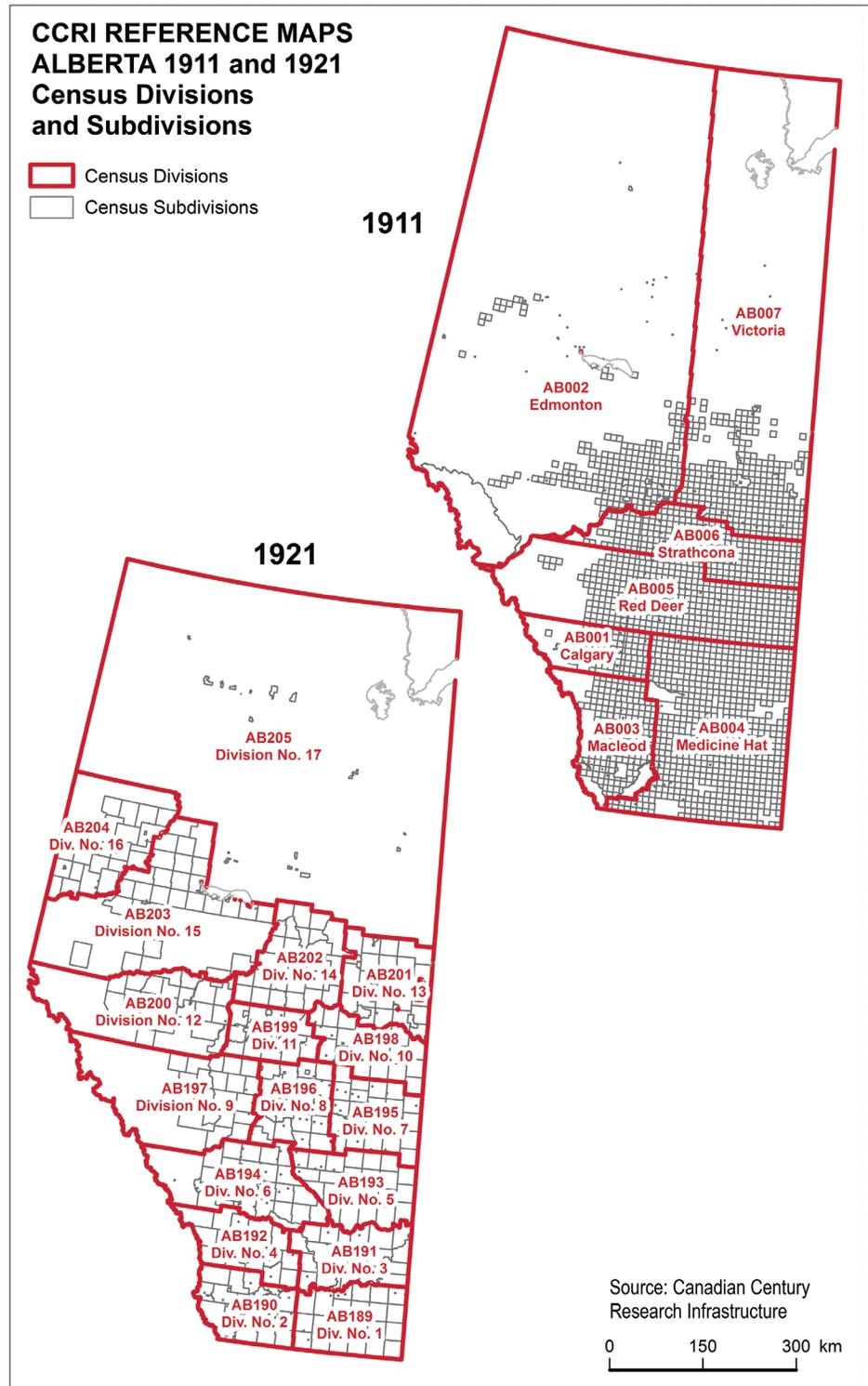
The Census of Canada may appear monolithic from afar, but in truth it is a dynamic instrument that has changed every time it has been conducted. These changes are not limited to changes in geographic boundaries and changes in the questions that are asked; they also include the political influences, the administrative organization, and the personnel and methods used to carry out the assembling of data intended to capture a snapshot of the people and activities of the nation. Researchers should be aware of these issues in working with the CCRI data.

In general, the officers of the census were instructed to use the federal electoral divisions and subdivisions at the time to create the enumeration areas for the collection of census data. These adapted electoral divisions were usually termed census *districts* and *subdistricts*. In contrast, the reporting of the census was to be done on the basis of municipal geographic units: typically counties, townships or parishes, cities, towns, and villages; these were usually termed census *divisions* and *subdivisions*.¹⁹ Some literature refers to the collection process as the “census-taking,” and the reporting (i.e., the publishing of the results of the census in official printed volumes and bulletins) as the “census-making.”

What we learned by delving into the census data and documentation of 1911–51 was how ephemeral this census-taking and census-making infrastructure could be. It was also poorly documented, especially in early years.²⁰ An advantage of having Statistics Canada as a partner in the CCRI project was that researchers were given access to their historical files. No maps, however, were found of census enumeration areas or even of census subdivision boundaries for the early part of our time period. Some maps appear in the published volumes from 1921 onward, but they vary in their level of detail and are frequently only available at the census-division level. By 1951 census tract maps had been published for the newly established Census Metropolitan Areas, but even in that year census subdivision boundaries in rural areas were still poorly defined.²¹

Also unexpected was the amount of deviation between the system used for census-taking and that used for census-making. From previous experience with census data, specifically 1901, it was assumed that the geographic areas

Fig. 13.1.
 Comparison between
 census divisions
 and subdivisions
 in Alberta in
 1911 and 1921.
 (Source: Canadian
 Century Research
 Infrastructure. Map
 produced by the GIS
 and Cartography
 Office, Department
 of Geography
 and Program
 in Planning,
 University of
 Toronto.)



used for these would be coincident, or at least hierarchically nested: individual enumerator's areas and enumeration subdistricts would be grouped together within published census subdivisions and divisions. This turned out to be emphatically not the case. The enumeration areas, subdistricts, and districts used to collect the census data were often significantly different from the subdivisions and divisions used to report the census data.²²

Another issue that should be understood by users of the CCRI data is the significant differences in the definition of census geography – for example, what comprises a census subdivision in different parts of the country and for censuses taken at different times. Generally, the earlier-settled agricultural areas of the country have a more intensive and persistent system of townships, towns, and villages comprising CSDs. The less-settled northern reaches have more extensive and inconsistent definitions. The most critical concerns about these differences arise when users wish to make comparisons over time. A look at the change in the numbers of census subdivisions between the different census periods puts the magnitude of these changes into perspective. For example, Figure 13.1 shows that in 1911 Alberta had 2,256 CSDs grouped into seven CDs; in 1921 in Alberta the number of CSDs was down to 520, but the number of CDs was up to seventeen.²³ Although the changes between 1921 and later census years were less dramatic, the units of census geography continued to shift and evolve.

Fortunately for users, these incompatibilities are no longer incapacitating. The CCRI geographic team has reconstructed the geographic boundaries of census subdivisions to cover the time period.²⁴ Using federal electoral

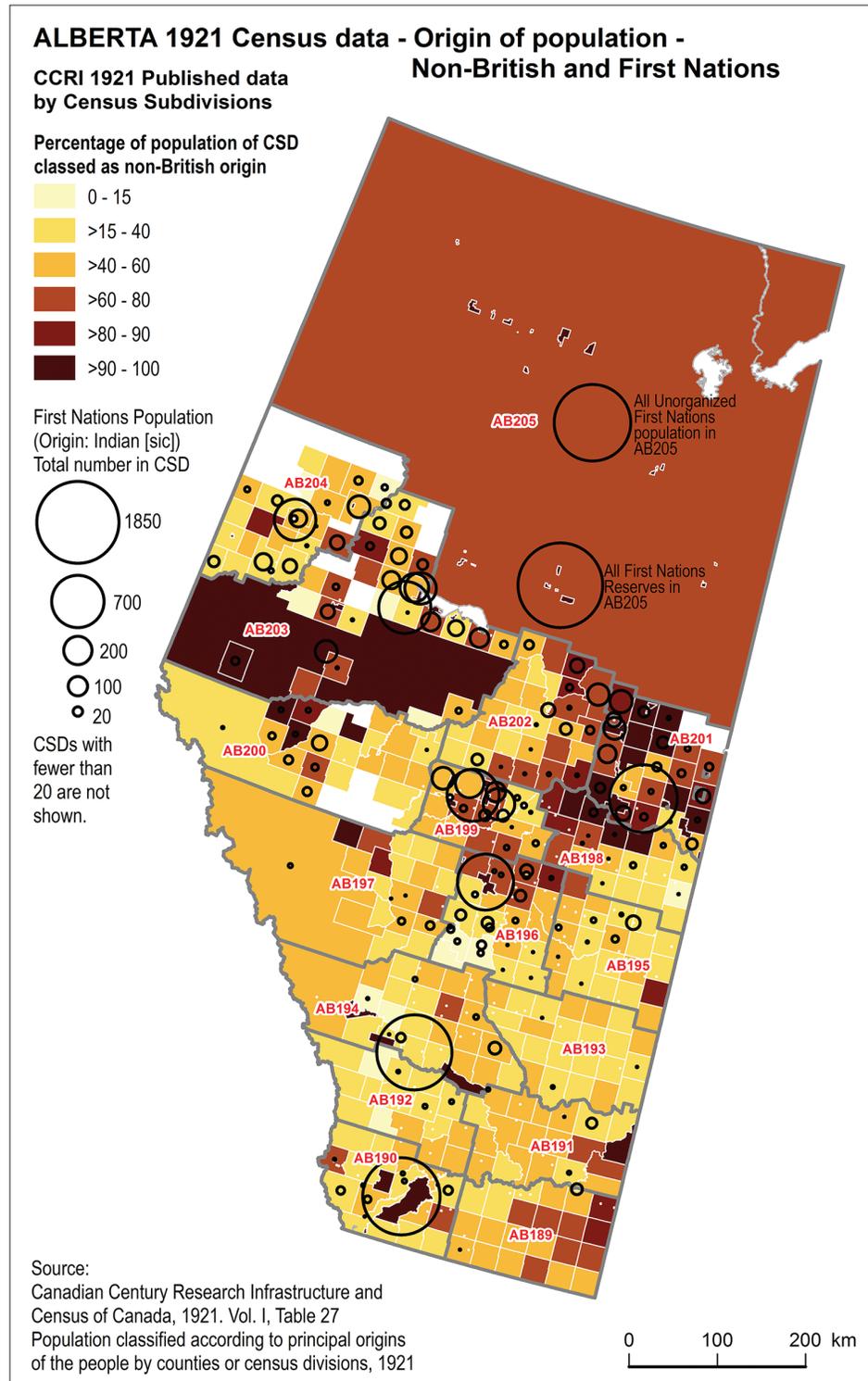
atlases published at regular intervals, textual descriptions of census subdivisions in the published census volumes, and contemporaneous historical maps from Library and Archives Canada and other sources, the team created polygons of the components of the CSDs for this period and linked them with the microdata records, allowing researchers to proceed with confidence that data will be matched successfully. The maps and GIS files resulting from this effort, furthermore, can be used to establish compatibility and to enable comparisons between censuses.

WORKING WITH THE CCRI DATA USING GIS

Historical GIS analysis using the CCRI data is similar to other uses of census data in a GIS environment, with a few additional burdens associated with the use of sample data (rather than aggregated data based on the full population).²⁵ Either selected and grouped microdata or aggregate data from the digitized published tables may be linked to the CCRI GIS files to be mapped or analyzed. It is expected that most users will be selecting and grouping microdata and doing thematic mapping of the results. Mapping the data allows researchers to visualize geographic distributions and patterns and compare these patterns between different variables, places, and times.

To offer a simple example, a researcher might wish to thematically map some of the data from the digitized published tables using the CCRI GIS files. The first step would be to examine the published tables and the documentation describing how they were created.²⁶

Fig. 13.2.
 Example of thematic map created by linking CCRI digitized published table to CCRI geographic file. (Source: Canadian Century Research Infrastructure and Census of Canada, 1921, vol 1, Table 27. Map produced by the GIS and Cartography Office, Department of Geography and Program in Planning, University of Toronto.)



These aggregate tables present data either at the census-subdivision level or at the census-division level. The CCRI reference maps provide a helpful supplementary reference for reviewing the census geography of the time and place in order to facilitate data selection and download. Both the GIS files and the digitized published tables have a table-specific identification field; for example, “V1T1_1911” is the ID field that appears in the table taken from Volume I, Table 1 of the 1911 published census. This identifier also appears as a field in the map file identifying the corresponding polygon. When the GIS file and the data table are brought into GIS software, the files can be joined using this common variable, and mapping of the aggregate data can proceed. Figure 13.2 presents such a map for Alberta, using the 1921 published data table of “principal origins” of the population, digitized by CCRI. It shows the percentage of the population of “non-British” origins by CSD; it then superimposes proportional circles showing numbers of First Nations people.

Working with the microdata is undeniably more complex. Coding and cleaning of these data has been done, and the documentation for this is available on the CCRI website. Since these tables are based on individual census records, however, all classification and aggregation must be done by the user. Typically, researchers create their own classifications based on their own research needs. For example, to establish a series of household classifications, social historian Gordon Darroch examined each head of household, and the relationship of the other members of the household to him/her (see Table 13.1).²⁷ This required querying a number of variables in the table – primarily the “Relationship to Head,” “Marital Status,” and “Gender” fields, for their appropriate codes

– and then creating a new variable to assign a “Household Situation” value. Darroch generated twenty-five of these values, which he then coalesced into eight main “household types,” and then combined further into three “household classes.” This work may be done interactively in a statistical package like SPSS and then used to generate scripts to document and replicate the process.

The CCRI geographic framework uses as its basic polygon unit the “CCRIUID” – a unique ten-character geographic identifier that forms the key link between the geographic files and the census microdata. It is an alphanumeric variable based on the census division and subdivision numbers as listed in a selected “key” published table. This hierarchical geographical coding system is modelled on Statistics Canada’s modern digital boundary file coding. Every record in the CCRI microdata file has been given a CCRIUID, corresponding to the GIS file. Once a researcher’s classification has been completed, the data can be aggregated by CCRIUID and then collapsed into any other level of census geography (census subdivision, division, province). The result is a table with, for example, numbers of individuals in households of each of the eight “household types,” within each census subdivision. In a GIS package, this table can then be joined to the GIS file of census subdivisions and manipulated to create a map or do spatial analysis, as required. Figure 13.3 illustrates thematic mapping of one adaptation of Darroch’s household classifications, which uses four “household classes,” developed using the CCRI sample microdata for 1911.

There are a few methodological issues to bear in mind when using the CCRI microdata. Sample size is the most significant issue. The CCRI aimed for a 5 per cent sample in

Table 13.1. Hierarchical classification of household situation based on “Relationship to head.”

Level A. Household Situation (25).					Level B. Household Type (8).		Level C. Household Class (3). (Darroch)	
		Living with...						
Code	Status of head	Kids*	Kin*	Non-kin*				
0.	Married but alone				1.	Primary individuals	1.	Lone head households
0.		no	no	no	2.	Single parents	2.	Married couple households
	Married				3.	All other lone head with non-kin	3.	Extended households
1.		yes	yes	yes	4.	Childless couples	Level C. Household Class (4). (revised for mapping)	
2.		yes	yes	no	5.	Couples with children	1.	Single individuals
3.		yes	no	yes	6.	All couples with non-kin	2.	All other single heads
4.		yes	no	no	7.	Extended household	3.	All couples (no kin)
5.		no	yes	yes	8.	Extended households with non-kin	4.	All extended (kin)
6.		no	yes	no				
7.		no	no	yes				
8.		no	no	no				
	Male Single							
9.		yes	yes	yes				
10.		yes	yes	no				
11.		yes	no	yes				
12.		yes	no	no				
13.		no	yes	yes				
14.		no	yes	no				
15.		no	no	yes				
16.		no	no	no				
	Female Single							
17.		yes	yes	yes				
18.		yes	yes	no				
19.		yes	no	yes				
20.		yes	no	no				
21.		no	yes	yes				
22.		no	yes	no				
23.		no	no	yes				
24.		no	no	no				

Source: Gordon Darroch (personal communication, e-mail, 28 April 2010).

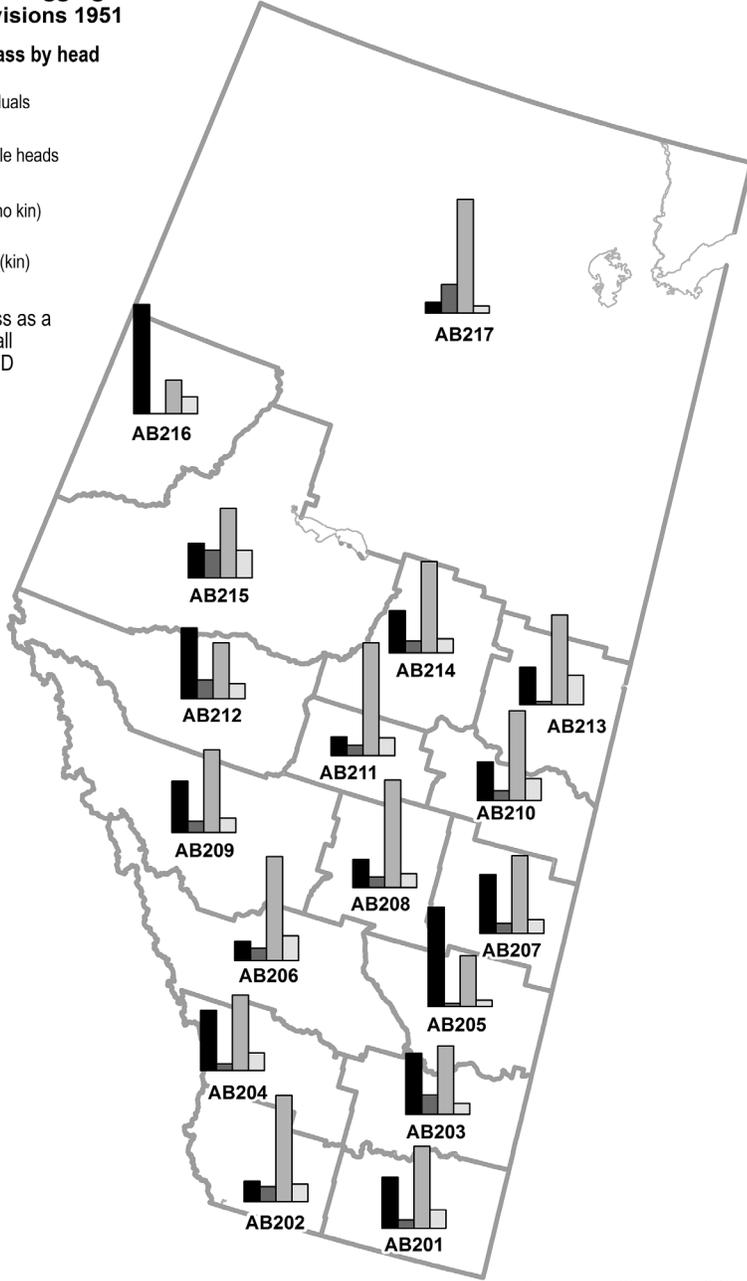
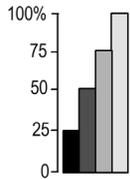
ALBERTA 1911 Census data - Household class by Census Divisions

CCRI 1911 Microdata sample weighted and aggregated to Census Divisions 1951

Household class by head

- Single individuals
- All other single heads
- All couples (no kin)
- All extended (kin)

Household class as a percentage of all individuals in CD



Source: Canadian Century Research Infrastructure

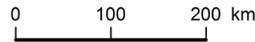


Fig. 13.3. Example of thematic map created by linking CCRI classified sample data for 1911, weighted and aggregated to the CCRI geographic file of 1951 Census divisions. (Source: Canadian Century Research Infrastructure. Map produced by the GIS and Cartography Office, Department of Geography and Program in Planning, University of Toronto.)

Table 13.2. Sampling rates, sample sizes, and population estimates for each CCRI census year.

Census year	Sampling rate of “Regular dwellings”	Sampling rates within “Large-dwellings” of single- or multi-unit type		Sample sizes (Ns)		CCRI weighted Canadian population estimate	Published Census population of Canada
		Single-units (by individual)	Multi-units (by unit)	(dwellings)	(individuals)		
1911	5%	10%	25%	73,141	371,373	7,164,970	7,206,643 (+0.6%)
1921	4%	10%	20%	73,505	367,475	8,707,780	8,788,483 (+0.9%)
1931	3%	10%	20%	65,331	348,519	10,277,240	10,376,786 (+0.9%)
1941	3%	10%	20%	80,131	363,935	11,381,887	11,506,655 (+1.1%)
1951	3%	10%	n/a	105,620	443,921	13,915,210	14,009,429 (+0.6%)

Source: Canadian Century Research Infrastructure User’s Guide and internal documentation, including Statistics Canada internal certification reports for inclusion of data in Research Data Centres.

1911, 4 per cent in 1921, and 3 per cent for the other three census years. Higher sampling rates were constructed for large dwellings (defined as those with more than thirty residents) to capture the institutional population, so these need to be weighted differently in creating overall population estimates. Sample sizes for all dwellings and the Canadian population are large enough to be adequate for most analyses (see Table 13.2). The standard statistical software packages allow population estimates to be made automatically based on varying sampling rates. Issues will arise, however, when data are extracted for geographic areas with small populations such as census subdivisions, and even for some census divisions. One of the strengths of census microdata is the ability to analyze subgroups of the population.²⁸ But when a subpopulation is the focus of research, such as a specific age/sex cohort or an occupational group, small sample sizes restrict the analytical capabilities of the data.

Another concern with small sample sizes relates to confidentiality issues for the 1921–51 period. Work may be done on these files at the individual level in the RDC, but results taken out, and especially published results, are subject to disclosure rules designed to maintain confidentiality. Data on individuals cannot be disclosed or mapped; it must first be cross-tabulated by geographic area, and table cell sizes of at least five individuals are required. As a general rule of thumb to ensure representativeness, CCRI guidelines suggest at least twenty sample dwellings or one hundred individuals per geographic unit. Since there are many instances when smaller sample sizes occur, researchers must aggregate cases by geographic areas. Geographic aggregations may be created by standard census geographic areas (e.g., more than one CSD or CD grouped together); by census attributes (e.g., contiguous regions of similar population density grouped together); or by externally defined geography

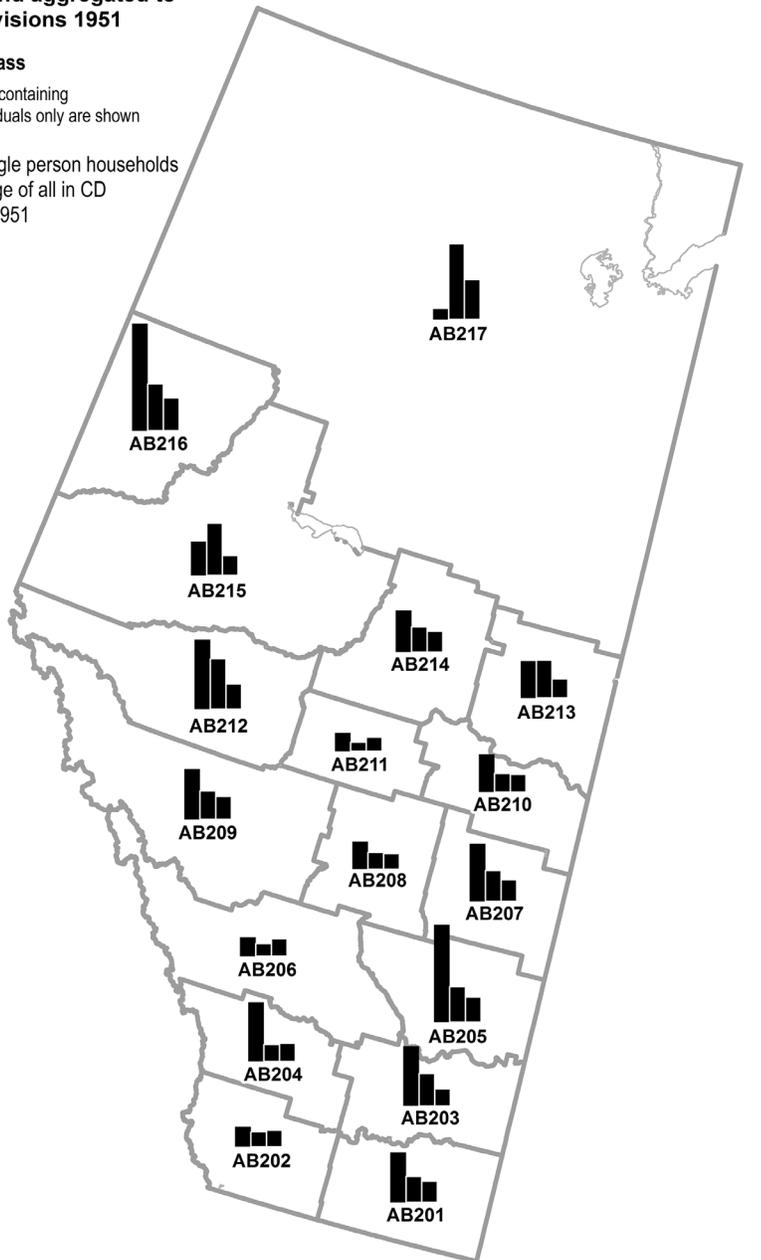
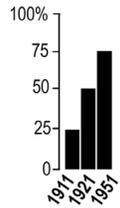
ALBERTA 1911-1921-1951 Census data - Single person households by Census Divisions

CCRI Microdata samples 1911-1921-1951
weighted and aggregated to
Census Divisions 1951

Household class

Households containing
Single individuals only are shown

Change in Single person households
as a percentage of all in CD
1911 - 1921 - 1951



Source: Canadian Century Research Infrastructure

Fig. 13.4.
Example of cross-census comparison map created by linking CCRI classified sample data for three census years, weighted and aggregated to the common base of 1951 Census divisions. (Source: Canadian Century Research Infrastructure. Map produced by the GIS and Cartography Office, Department of Geography and Program in Planning, University of Toronto.)

(areas of similar land characteristics such as urban or rural areas, or ecological zones, may be grouped together).

The decision about how to aggregate will depend upon the framing of the research question, and the amount of aggregation necessary to meet Statistics Canada's disclosure requirements. An example of this is shown in Figure 13.4, which illustrates the change in the proportion of single-person households in Alberta between 1911 and 1951. In order to do this comparison, the classified households for 1911, 1921, and 1951 were aggregated by the census divisions of 1951. Because there were relatively large numbers in the sample for single-person households, the 1921 and 1951 data were approved for disclosure. Other household types such as "single parents" had relatively small numbers; researchers working with these data would be required to combine some of the 1951 CDs in order to meet the Statistics Canada disclosure requirements.

CONCLUSION

Research using the CCRI is a developing story. Since the data are still in the process of being made available to users, this introduction to the CCRI can only hint at its potential. Similar projects have experienced distinct alternating phases of resource creation and methodological innovation.²⁹ The CCRI is entering such an innovation stage. It invites interested users to follow their own lights in utilizing this resource and to add their voices to the conversation.

The next steps for the CCRI will improve access to the data and other tools. In the long term, online open access to aggregated

anonymized data is the goal. In addition to simple access, building tools for online geographic selection, aggregation, and mapping of results will enhance the usefulness of the geographic component of the data. Several additional projects would add considerably to functionality of the CCRI data. One would be to establish cross-census harmonization methods and best practices for these and other historic Canadian census data. Data expansion and enhancement is also a realistic goal; the CCRI geographic data were designed to be extensible and capable of infill. One important objective that could take advantage of other historical GIS work done elsewhere would be to fill in the census geography and supplement microdata geocoding at the intra-urban level for major cities.

The geographic team on the CCRI has always seen the project as part of the larger picture of conducting and enhancing spatial historical studies in Canada. As an infrastructure project, it is only one piece in what should be a shared resource of historical data and mapping made available to scholars and the public. This may be conceived of as part of a "national historical GIS" comparable to those in the United States and Britain. The experience of the CCRI indicates that academics, government, and the private sector can work together to develop such resources. As Gregory and Ell commented on the Great Britain HGIS project, "Although the investment that has been put into this may seem large, compared to the investment in collecting the statistics that they contain, these resources are small, and the potential for reinvigorating [their] use ... is large."³⁰ Building a partnership to leverage this investment seems like a good place to start.

NOTES

- 1 Library and Archives Canada holds the early manuscript census schedules on microfilm or fiche and is in the process of scanning them and making them available online, either through their own website or those of genealogical societies (see <http://www.collectionscanada.gc.ca/genealogy/022-911.009-e.html>). The Automated Genealogy project (<http://automatedgenealogy.com/>) has made digital images for several censuses available and indexes them by name. Statistics Canada holds all census records for 1921 up to the present securely in its archives.
- 2 The University of Ottawa was the lead institution for the CCRI project. Other university partners included the Université du Québec à Trois Rivières (Centre interuniversitaire d'études québécoises – CIÉQ); Université Laval (CIÉQ); Memorial University of Newfoundland; University of Toronto; York University; University of Victoria; and the University of Alberta, which continues to host the CCRI archive and distribute its microdata for 1911. Institutional and government partners included: Statistics Canada; Newfoundland and Labrador Statistics Agency; Institut de la statistique Québec; International Microdata Access Group; and Library and Archives Canada. The CCRI was supported by the Canada Foundation for Innovation, the Ontario Innovation Trust, the FCAR Funds (Quebec), and the Harold Crabtree Foundation.
- 3 Chad Gaffield, "Introduction. Conceptualizing and Constructing the Canadian Century Research Infrastructure," *Historical Methods* 40, no. 2, Special Issue on the Canadian Century Research Infrastructure (2007): 55.
- 4 See <http://ccri.library.ualberta.ca/>.
- 5 Gordon Darroch, Richard D. B. Smith, and Michel Gaudreault, "CCRI Sample Designs and Sample Point Identification, Data Entry, and Reporting (SPIDER) Software," *Historical Methods* 40, no. 2, Special Issue on the Canadian Century Research Infrastructure (2007): 65–75.
- 6 Chad Gaffield, "Evidence of What? Changing Answers to the Question of Historical Sources as Illustrated by Research Using the Census," in *Building New Bridges: Sources, Methods and Interdisciplinarity*, ed. Jeff Keshen and Sylvie Perrier (Ottawa: University of Ottawa Press, 2005), 265–74.
- 7 C. Bellavance, F. Normand, and E. S. Ruppert, "Census in Context: Documenting and Understanding the Making of Early-Twentieth-Century Canadian Censuses," *Historical Methods* 40, no. 2, Special Issue on the Canadian Century Research Infrastructure (2007): 92.
- 8 See <http://ccri-cd.cieq.ca>, also accessible through <http://ircs1911.cieq.ca>.
- 9 Similarities exist to earlier projects such as the Great Britain Historical GIS project (<http://www.gbhgis.org>) and to contemporaneous efforts such as the American National Historical Geographic Information System project (<http://www.nhgis.org>).
- 10 See <http://www.ipums.org>.
- 11 See <http://www.nappdata.org>.
- 12 See, respectively, the Programme de recherche en démographie historique (<http://www.prdh.umontreal.ca/census>) and the Canadian Families project (<http://web.uvic.ca/hrd/cfp/>). Two publications based on the Canadian Families project were particularly important to the CCRI project and the research examples in this chapter: a special issue on the Canadian Families Project in the Fall 2000 issue of *Historical Methods* (33, no. 4); and Larry McCann, Ian Buck, and Ole Heggen, "Family Geographies: A National Perspective," in *Household Counts: Canadian Households and Families in 1901*, ed. Eric Sager and Peter Baskerville (Toronto: University of Toronto Press, 2007).
- 13 See the 1891 Census of Canada website (<http://www.census1891.ca>) and the BALSAC website (<http://wordpress.uqac.ca/~balsac-en/>) respectively.
- 14 Gaffield, "Introduction," 57.
- 15 The NESSTAR data download service for the CCRI database for 1911 is accessible at: <http://nesstar2.library.ualberta.ca/webview/>.
- 16 See the Statistics Canada Research Data Centres website at: <http://www.statcan.gc.ca/rdc-cdr/index-eng.htm>.

- 17 See the CCRI Geographic Files Access website at: <http://ccri.library.ualberta.ca/endatabase/geography/gislayers>.
- 18 See the Statistics Canada Dissemination Area files documentation at: <http://www.statcan.gc.ca/bsolc/olc-cel/olc-cel?lang=eng&catno=92F0169X>.
- 19 According to the instructions to the Officers of the 1911 Census, and the Census Act of Canada (Revised Statutes of Canada, 1906, c. 68), *census districts* “shall correspond as nearly as may be with the existing electoral divisions and subdivisions; and the said Census districts may be further divided into such sub-districts as the Minister of Agriculture may direct as units of enumeration for Census purposes.” Census sub-districts “shall ordinarily consist of townships, parishes, cities, towns and incorporated villages.” In the published census volumes the use of the terms “division” and “district” are used differently in different years, as summarized below:
- 1911: Census districts and subdistricts used for all published tables; Electoral divisions and subdivisions used in instructions to Officers.
- 1921: Census divisions and subdivisions used for the majority of published tables; Electoral districts and subdistricts used for some published tables. Census districts and subdistricts referred to for enumeration purposes in instructions to officers.
- 1931: Census divisions and “municipality, township or subdivision” used for all published tables.
- 1941: Census divisions and subdivisions used for the majority of published tables; Electoral districts used for some published tables.
- 1951: Census divisions and subdivisions used for the majority of published tables; Electoral districts used for some published tables.
- 20 Byron Moldofsky, “The CCRI Geographic Files: Introduction and Examples,” in *The Dawn of Canada’s Century: Hidden Histories*, ed. Gordon Darroch (Montreal and Kingston: McGill-Queen’s University Press, 2013) (forthcoming).
- 21 Canada. Department of Trade and Commerce, Dominion Bureau of Statistics, *Ninth Census of Canada, 1951* (Ottawa: Queen’s Printer, 1953), vol. 1, Appendix A, “Reference Maps.”
- 22 Marc St-Hilaire et al., “Geocoding and Mapping Historical Census Data: The Geographical Component of the Canadian Century Research Infrastructure,” *Historical Methods* 40, no. 2, Special Issue on the Canadian Century Research Infrastructure (2007): 80.
- 23 *Ibid.*, 78.
- 24 St-Hilaire et al., “Geocoding and Mapping Historical Census Data”.
- 25 A discussion of CCRI sample designs and their implications is included in Darroch, Smith, and Gaudreault, “CCRI Sample Designs.” A further explanation is available in the online user’s guide at <http://ccri.library.ualberta.ca/endatabase/sampling/sampledesign/>.
- 26 The digitized published tables are available in the online user’s guide at <http://ccri.library.ualberta.ca/endatabase/geography/digitizedpublictables/index.html>.
- 27 Gordon Darroch, personal e-mail communication, 28 Apr 2010. Darroch analyzes numbers of individuals in different household situations rather than numbers of households themselves. So do the examples here. See also Gordon Darroch, “Household Experiences in Canada’s Early 20th Century Transformation,” in *The Dawn of Canada’s Century: Hidden Histories*, ed. Gordon Darroch (Montreal and Kingston: McGill-Queen’s University Press, 2013) (forthcoming).
- 28 Angela Dale, Ed Fieldhouse, and Clare Holdsworth, *Analyzing Census Microdata* (New York: Oxford University Press, 2000), 30.
- 29 Ian N. Gregory and Paul S. Ell, *Historical GIS: Technologies, Methodologies, and Scholarship* (Cambridge: Cambridge University Press, 2008), 189.
- 30 *Ibid.*