

Canadian Census Data Discovery Partnership (CCDDP)
<https://cddp-pddr.ca>

CCDDP CONSULTATION REPORT #2

Disseminating Census Data: Learning from Experience

VIRTUAL WORKSHOP

February 16, 2022

Zack Taylor, Western University
Susan Mowers, University of Ottawa
Shannon Leitch, University of Ottawa
Julien Doris, Université d'Ottawa
Leanne Trimble, University of Toronto

April 21, 2022

Également disponible en français :

Diffuser les données du recensement : L'apprentissage par l'expérience

Contents

Summary of Lessons	3
Audience, Usability, Discoverability, and Access	3
Long-term Sustainability of Projects	3
Frontiers of Innovation	3
Data Privacy	3
Introduction	3
Overview	4
Survey of presenters	5
Theme 1: Audience, Usability, Discoverability, and Access	7
Defining the audience	7
Access-usability-cost tradeoffs	9
Maximizing discoverability and access (for the right audience)	9
Lessons	10
Theme 2: Long-term Sustainability of Projects	10
Financial sustainability	10
Technological sustainability	11
Planning for the inevitable: Archiving and preservation strategies	11
Sustainable data stewardship	12
Lessons	12
Theme 3: Frontiers of Innovation	12
Advances in on-line analysis and visualization tools	12
New data access tools	13
Value-added products	13
Digitization	13
Microdata linkage	13
Lessons	14
Theme 4: Data Privacy	14
Changing privacy norms and requirements	14
Toward flexibility? The access/privacy tradeoff	14
Disclosure control techniques	14
Lessons	15
Appendix A: Presenters and Projects	16
Session 1 – Challenges and Innovations	16
Session 2 – Disseminating Microdata: International Perspectives	19
Session 3 – Disseminating Microdata: Canadian Perspectives	21
Appendix B: Overview of Census Discovery Platform Features as Demonstrated	25

Summary of Lessons

Audience, Usability, Discoverability, and Access

1. Determining a census dissemination project's audience(s) is an essential early decision. Important decisions about the design of the dissemination channel cannot be made without clarity on the identity and needs of user groups.
2. To meet the needs of different user groups, and in an environment of scarce resources, projects must carefully balance tradeoffs between access, usability, and cost. This requires consideration of the most likely use cases.

Long-term Sustainability of Projects

3. Data and purpose-built dissemination portals created by university-based researchers using time-limited grants are vulnerable to funding loss, personnel turnover, and changes in institutional priorities.¹
4. Projects should include an archiving and preservation strategy from the start, including the identification of a responsible party to steward the data and associated legal obligations.

Frontiers of Innovation

5. Accelerating computing power and storage have unlocked a range of data processing and dissemination innovations.
6. As the pace of innovation increases, the twin challenges of fiscal and technological sustainability will only increase.
7. Technological change means analysis of microdata will become more feasible for researchers in academe, government, and the private sector, posing challenges for data security and privacy.

Data Privacy

8. Social norms and legal requirements regarding data privacy are in flux, largely in the direction of stronger protection.
9. It is unclear what long-term impact this will have on research use of census and other data.
10. Different risk-reward trade-offs regarding the tension between access and privacy are possible.

¹ Alison Paprica. (2021, March 14). *Risks for academic research projects, an empirical study of perceived negative risks and possible responses* (Scholarly report no. arXiv:2103.08048). Cornell University. <https://doi.org/10.48550/arXiv.2103.08048>

Introduction

Overview

Censuses, or population counts, have been conducted in the territory now known as Canada since 1665–66 in New France. The census is Canada’s most valuable primary economic, social, and cultural dataset, and is an essential research tool for the formation of new knowledge and understanding about the populations that lived here in the past and present. The **Canadian Census Data Discovery Project** is a two-year project to improve access to and research conditions for existing qualitative, quantitative and spatial data sources from the population censuses of Canada. The key deliverables will be the design and delivery of a proof-of-concept bilingual discovery portal, and future-looking recommendations upon which this project can be further built.

To these ends, the Project has conducted a series of stakeholder consultations. In collaboration with Western University’s **Network for Economic and Social Trends**, the Project held the second of these on February 16, 2022, on the topic of ***Disseminating Census Data: Learning from Experience***. The invitation-only workshop’s focus was on challenges associated with the dissemination of contemporary and historical census data to academic researchers and the public. The goal was to better understand how creators and stewards of Census data, inside and outside of government, approach dilemmas associated with its dissemination to stakeholders, and also to create an international community of practice regarding these matters.

As our goal was to obtain a broad range of perspectives, we invited presenters from government statistics agencies and academic Census data projects in Canada, Ireland, the United Kingdom, and the United States. As our focus is on open access in the public interest, we did not invite for-profit data vendors to the workshop.

The workshop unfolded in three thematic sessions moderated by Project investigators. The presentations by the 14 presenters were followed by questions from invited members of the government statistics stakeholder community.

Session 1: Challenges and Innovations

- Jonathan Schroeder, IPUMS–National Historical GIS, USA
- Humphrey Southall & Paula Aucott, A Vision of Britain Through Time Project, UK
- Glen Hohlmann & David Price, Statistics Canada
- Elaine Castonguay, Statistics Canada

Session 2: Disseminating Microdata, International Perspectives

- Deirdre Lynch, Central Statistics Office, Ireland
- Callum Foster, Office for National Statistics, UK
- Lara Cleveland, IPUMS, USA

Session 3: Disseminating Microdata, Canadian Perspectives

- Lexi Clemence, Libraries and Archives Canada
- Sandra Sawchuk, Mount St. Vincent University
- Susan Mowers, University of Ottawa
- Kris Inwood, The Canadian Peoples Project
- Michael Haan, Western University

This consultation report contains a thematic summary of the workshop presentations and draws lessons for future Census dissemination initiatives. The presentations are hosted on the project website.

Survey of presenters

We surveyed the presenters regarding their projects' origins, goals, and funding arrangements prior to the workshop. This elicited useful responses, some of which are summarized in Tables 1–3. The full responses to the survey, and biographical information about the presenters and their projects, are reproduced in the appendix.

Table 1: Project Origin

	As an academic research project	As a government initiative
IPUMS–NHGIS	✓	
A Vision of Britain Through Time (UK)	✓	
Census Dissemination (StatCan)		✓
Census Geography (StatCan)		✓
IPUMS–International	✓	
Central Statistics Office (CSO-Ireland)	✓	
Office for National Statistics (ONS-UK)		✓
Library and Archives Canada (LAC)		✓
Canadian Century Research Infrastructure (CCRI)	✓	
Canadian Census Data Discovery Partnership (CCDDP)	✓	
The Canadian Peoples Project (TCP)	✓	
Research Data Centres (RDC-Canada)	✓	

While the reported origin of the projects (Table 1) was no surprise – by and large, government projects were initiated by governments and university-based projects were initiated by academic researchers – we found significant variation in the projects' goals (Table 2). While virtually all of the projects prioritized sustained access to users, they made different tradeoffs regarding the preservation of original formats versus the processing or reformatting of the original data to facilitate topic or geographic linkage across time. Several of the government and

university-based projects expressly valued innovation, while others did not. Government agencies in particular prioritized regular reporting. This is unsurprising given their accountability structure. Only a few prioritized partnerships with other public and private entities.

Table 2: Project Goals

	Sustained access	Open access	Preservation of original formats	Cross-time linking	Cross-time harmonization	Innovation	Regular reporting
IPUMS–NHGIS	✓	✓		✓	✓	✓	
A Vision of Britain	✓			✓	✓		
Census Diss. (StatCan)	✓	✓	✓				✓
Census Geography (StatCan)	✓	✓		✓	✓		
IPUMS–Int'l	✓			✓	✓	✓	
CSO-Ireland			✓		✓		✓
ONS-UK	✓	✓	✓	✓	✓	✓	✓
LAC	✓		✓				
CCRI	✓	✓				✓	
CCDDP	✓	✓			✓		
TCP	✓			✓	✓	✓	
Research Data Centres (RDCs Canada)	✓		✓	✓	✓	✓	✓

The respondents also reported a range of funding sources for their projects (Table 3). Government statistics collection and dissemination services depend on government program funding, while university-based projects generally launch with support from governments, typically via competitive academic research support grants, and sometimes also with internal support from universities and external support from foundations and charities. As startup funds are used up, university-based projects turn to other sources of funding, including from foundations and charities, and in some cases from licensing and advertising. As discussed in greater detail below, the survey revealed significant concerns about the fiscal sustainability of university-based projects and ongoing maintenance of project infrastructure following their completion.

Table 3: Project Funding Sources

	Government program funding	Grants from governments, including academic research support	Non-profit foundations or charities	Universities	Cost-recovery from paying users, licensing, web ads
IPUMS–NHGIS		Primary	✓	✓	Donations
A Vision of Britain		Startup	Startup		Current
Census Diss. (StatCan)	✓				
Census Geography (StatCan)	✓				
IPUMS–International		Primary	✓		
CSO-Ireland	✓				
ONS-UK	✓				
LAC	✓				
CCRI		✓			
CCDDP		✓			
TCP		✓			
RDCs-Canada	✓				

Theme 1: Audience, Usability, Discoverability, and Access

The first theme that emerged from the workshop discussion was the relationship between three factors:

- *usability* (the appropriateness of tool complexity for the intended audience),
- *discoverability* (the ability for users to find what they are looking for, or to find the dissemination channel in the first place), and
- *access* (the availability of data to users).

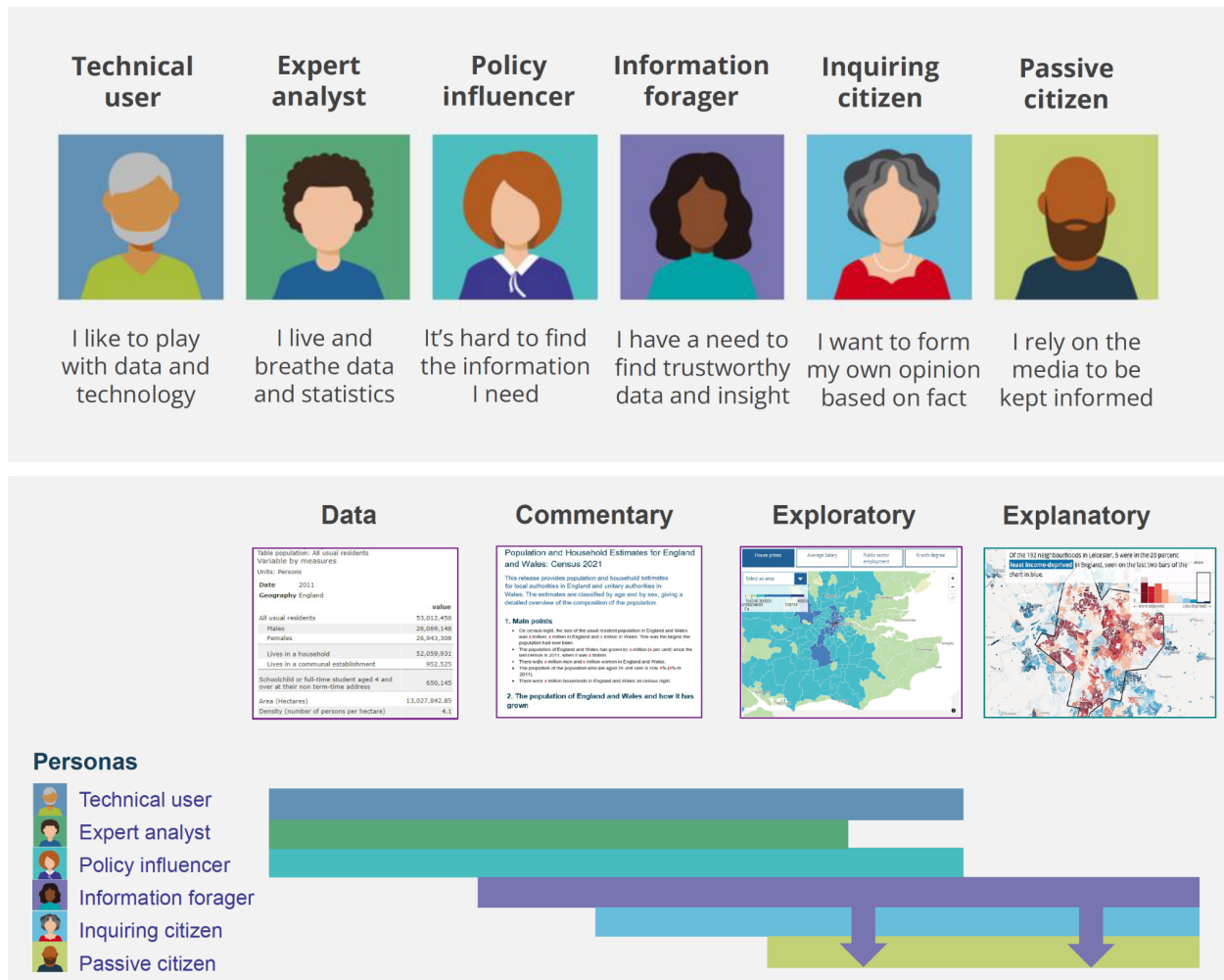
These factors hinge on the intended *audience* of the dissemination project. A project designed for sophisticated specialist users will embody different choices than one designed for the general public or casual professional users such as journalists.

Defining the audience

In his presentation, Callum Foster of the UK Office for National Statistics provided the most systematic understanding of project audiences. He outlined how the ONS’s dissemination

strategies and on-line tool design is driven by the needs of distinct user groups. Figure 1 arrays user *personas* from most to least sophisticated and identifies suitable products for them. While more sophisticated users (technical users, and expert analysts) want to explore and interact with the data directly, inquiring citizens and passive citizens are interested in receiving immediate insights and brief explanatory information, including visuals, for finding data. Policy influencers and information foragers lie somewhere between.

Figure 1: Different products for different user personas



Source: Foster, slides 2, 4.²

The other presenters were also cognizant of their project's intended audience, generally focusing on sophisticated users: researchers and analysts in government and universities, as well as in the nonprofit sector, and business. For example, the NHGIS and IPUMS facilities cater to sophisticated users who want to manipulate the data directly, so the output is a curated data

² Callum Foster. (2022, February 16). *Thinking about Census users and products, and ONS's Census 2021 dissemination and innovation* [Panel presentation]. Disseminating Census data: Learning from experience, a virtual workshop.

<https://cddp-pddr.ca/en/committees/user-needs/disseminating-census-data-event>

download rather than interactive analysis and visualization tools, with access to both longitudinal and cross-sectional data. (Jonathan Schroeder contrasted NHGIS with the for-profit Social Explorer web application, with which it shares data. Social Explorer includes user-queried mapping and visualization tools.) Similarly, the outputs of the Canadian Century Research Infrastructure and The Canadian Peoples projects are data products intended to be manipulated and analyzed by users themselves.

Unlike some of the other projects discussed, A Vision of Britain Through Time was expressly designed for use by the general public, without requiring any technical skills. Indeed, Humphrey Southall noted that they ran user needs and usability testing focus groups at the outset of their project. Their audience is ordinary people interested in particular *locations*; hence its focus on longitudinal rather than cross-sectional data access. Moreover, the public-facing element of Library and Archives Canada's census holdings primarily responds to requests about *people* – amateur and professional genealogists and historians.

Government statistics agencies must appeal to a broad range of potential audiences. They disseminate information about specific places to unskilled users through on-line portals, analysis and commentary in the form of reports, and also complete data sets for analysis by users, sometimes in secure settings, as with microdata within Canada's Research Data Centre network and equivalent facilities in the other countries.

Access-usability-cost tradeoffs

Foster's typology of audience personas and products illustrates the audience-driven tradeoffs between making data accessible to a broad audience and the complexity of the dissemination channel, and also between access and cost. Making complete data tables available for download (or queryable by API) is relatively inexpensive, but requires the user to be skilled. Web-based exploration, visualization, and analysis tools are expensive to construct and maintain, but may be designed to enable use by a wider range of people. In general, the more constrained the set of interaction choices available to the user, or the more curated rather than comprehensive the data made available, the easier it is for users to discover and retrieve data. At the same time, these constraints may make the data less useful to sophisticated users.

Maximizing discoverability and access (for the right audience)

Data are only usable if they can be discovered by their intended audiences. As official providers, government agencies can take discovery for granted. Glen Hohlmann noted that Statistics Canada is transitioning its public-facing Census data to the Government of Canada's new, centralized Corporate Output Data Repository. This may facilitate discoverability and access to diverse user groups.

Projects catering to specialist and expert users can appeal to user groups directly. Jonathan Schroeder and Lara Cleveland noted that the team that administers the NHGIS and IPUMS-International projects spends considerable time promoting them at scholarly and

professional meetings and events. By contrast, the CCRI project focused almost entirely on data digitization and validation with limited attention given to ongoing discoverability.

Humphrey Southall noted an important element of discovery, especially when the intended audience is ordinary people: ensuring that data pertaining to particular places and concepts appears in web searches, even if the pages are dynamically generated by database queries. The Vision of Britain portal was designed to expose web pages pertaining to particular places to search engines.

Finally, and connecting to the second theme, there is the matter of open access. All of the projects described disseminate their data to users at no cost unless they are restricted by privacy legislation or licensing terms. Nevertheless, pressure to recover costs through licensing or user fees increases as other sources of funding become scarce.

Lessons

1. Determining a census dissemination project's audience(s) is an essential early decision. Important decisions about the design of the dissemination channel cannot be made without clarity on the identity and needs of user groups.
2. To meet the needs of different user groups, and in an environment of scarce resources, projects must carefully balance tradeoffs between access, usability, and cost. This requires consideration of the most likely use cases.

Theme 2: Long-term Sustainability of Projects

A second theme that crosscut the panel discussions was the long-term sustainability of projects. This has at least three dimensions:

- *financial sustainability* (where will the money come from?),
- *technological sustainability* (who will maintain and update hardware and software?), and
- *sustainable data stewardship* (who will monitor time-delimited rights and licensing requirements, and also changing data privacy regimes?).

Financial sustainability

Most projects are worried about resources. Ongoing university-based projects are very concerned about financial sustainability. This is because the standard academic research model is to fund discrete, closed-ended projects with time-limited grants. Unless universities, foundations, or charities assume responsibility for funding projects on an open-ended basis, projects must turn to monetization strategies, including charging for desirable data products and analysis work, thereby undermining open access.

While benefiting from program funding, government statistics agencies face fiscal and political pressures of their own. The Census is expensive; there is always pressure to cut costs. As seen in the United States, and in Canada a decade ago, there is also political pressure from some quarters to make government data collection less intrusive by making surveys voluntary and eliminating questions deemed overly personal in nature. Agencies have responded as best they can by emphasizing continuity and increasingly relying on linked administrative data.

Ultimately, nongovernmental, university-based projects have different objectives than government programs. Government statistics agencies are present- and future-looking; their remit is to disseminate current data to policymakers and private-sector stakeholders in a timely manner. They have little incentive or interest in expanding, maintaining, and disseminating historical datasets by digitizing printed material, storing legacy data in modern file formats, or attaching modern geographic and concept identifiers to historical data. It falls to academic projects to perform these tasks, but with much more limited resources.

Technological sustainability

Presenters associated with university-based projects highlighted the problem of keeping up with technological change. One aspect of this is ensuring that the hardware and software – sometimes off-the-shelf but often custom-written – that underlie the dissemination channel are maintained and updated. Related to this is the problem of updating the data files to reflect changing data file formats and metadata standards. These tasks require specialized expertise and are costly to perform, hence technical and financial sustainability are inextricable.

Planning for the inevitable: Archiving and preservation strategies

The discussion turned to the question of what to do if an ongoing project is forced to end due to loss of financial resources or leadership turnover. This problem is especially acute for the Vision of Britain project. Humphrey Southall frankly acknowledged that the software and database management infrastructure that drive the website is failing and that there are no resources available to update it or modernize the web interface. As a fallback plan, he is pursuing a strategy of archiving flat tables on the UK Data Service. While this entails abandoning their innovative relational database model and undermines discoverability and access by their primary audience – ordinary people – it means that the data will be securely stored by a trusted and well-resourced steward.

This cautionary tale suggests the importance of a long-term archiving and preservation strategy for any data creation or reformatting project. Barriers to permanent archiving and preservation may be decreasing with the emergence of well-established data repositories with permanent base funding, such as Scholars GeoPortal, implementations of Harvard's Dataverse, the UK Data Service, and ICPSR. Some of these facilities are format-neutral; that is, it is possible to contribute and retrieve hosted data in a variety of formats, partially solving the problem of file format obsolescence. However, since these facilities generally store tables or boundary files

without selection and filtering tools, for example by location or variables, they are mainly of use to specialists.

Sustainable data stewardship

A long-term archiving and preservation strategy may ensure that data are not “lost” if their creators disappear. An additional issue is maintaining compliance with time-based access rules and licensing rules negotiated with private data vendors. For example, it was noted that the CCRI 1931, 1941, and 1951 microdata files (which are currently covered under the Government of Canada’s 92-year confidentiality rule), will enter the public domain in concert with the expiry of the 92-year period (e.g., 2023 for the 1931 CCRI microdata). but it is unclear whose job it is to ensure that embargoed files are transferred to open repositories. In addition, Kris Inwood described how the The Canadian Peoples project’s 100% microdata files were produced in partnership with a private vendor whose licensing terms preclude certain uses. Again, it is unclear who will be responsible for the data files over the long term.

Lessons

3. Data and purpose-built dissemination portals created by university-based researchers using time-limited grants are vulnerable to funding loss, personnel turnover, and changes in institutional priorities.
4. Projects should include an archiving and preservation strategy from the start, including the identification of a responsible party to steward the data and associated legal obligations.

Theme 3: Frontiers of Innovation

A key objective of the workshop was to identify frontiers of innovation. Synthesizing across the panel discussions, we identified five important developments:

- advances in on-line analysis and visualization tools,
- new data access tools (APIs),
- the creation of value-added products,
- the digitization of paper records, and
- microdata linkage.

These have been facilitated by increases in computing power and the emergence of fast and inexpensive cloud storage.

Advances in on-line analysis and visualization tools

Callum Foster’s demonstration of the ONS web application prototype, capable of tabulating microdata while applying disclosure control criteria in real time, was a highlight of the workshop.

Powered by cloud database storage and optimized algorithms, these tools blur the boundary between analyses performed on fast local computers and the convenience of web applications. They also blur the boundary between aggregate data and microdata. The days of costly special tabulations to create multidimensional cross-tabulations may be numbered.

New data access tools

There was also discussion of facilitating customized data retrieval by specialists through the use of application programming interfaces (APIs) – in essence, these are code libraries for software commonly used in data science and analytics that can programmatically query remote databases without manually retrieving data through a web-based graphic user interface. The NHGIS and IPUMS-International are developing APIs for direct data extraction into modern statistics software.

Value-added products

A third innovation is the development of standardized coding schemes for variables and geographic units, and also crosswalks and bridging tools that link geographic units across time. IPUMS-NHGIS has been a leader in this work, creating tables that apportion historical data to contemporary units. IPUMS-International has also put considerable work into developing a comprehensive data dictionary and metadata tools that enable cross-national comparison.

Digitization

Multiple Canadian university-based projects (CCRI and TCP) and LAC are engaged in the digitization of historical aggregate data and microdata. This has slowly filled gaps in the availability of digitized data. The Vision of Britain project and the NHGIS have done similar work in the UK and the United States, respectively. The ONS and the Irish Central Statistics Office have also participated in digitization of historical aggregate and microdata. Kris Inwood noted that as full-count historical microdata become available, it will become possible to cross-tabulate variables not published in the original printed volumes.

Microdata linkage

As computation power grows, it is increasingly possible to link administrative data to census data at the individual, household, or property level, or to link census data to other survey datasets (e.g., those pertaining to health, immigration, etc.). Michael Haan noted the growing use of Statistics Canada's restricted linked datasets pertaining to health outcomes and immigrant and refugee settlement. Deirdre Lynch discussed how Irish Census data have been linked to property and utilities data for use in long-term land-use and infrastructure planning. Glenn Hohlmann also mentioned how individual and household income data disseminated in Canadian Census products is now drawn from tax filer data rather than self-reported. As microdata linkage becomes more feasible, data stewards and researchers confront important questions regarding privacy and ethical use of data – the topic of the fourth and final theme.

Lessons

5. Accelerating computing power and storage have unlocked a range of data processing and dissemination innovations.
6. As the pace of innovation increases, the twin challenges of fiscal and technological sustainability will only increase.
7. Technological change means analysis of microdata will become more feasible for researchers in academe, government, and the private sector, posing challenges for data security and privacy.

Theme 4: Data Privacy

The final theme is data privacy. Participants identified three issues:

- *changing privacy norms and requirements,*
- *greater flexibility in access to microdata, and*
- *new disclosure control techniques applied to aggregate data.*

Changing privacy norms and requirements

The participants recognized that we live in an era of evolving attitudes toward privacy. On the one hand, “big tech” companies continually collect and monetize data about ordinary people’s movements and consumption habits. On the other hand, people are increasingly suspicious of governments data collection efforts. From the GDPR (General Data Protection Regulation) to new Canadian Treasury Board rules, legal and regulatory frameworks are evolving by requiring proactive permission from users and disclosure of how data will be used. Michael Haan asked whether we are being too absolutist as we seek to avoid identification of individuals, suggesting that we need to think about the likely practical risks of doing so in academic research.

Toward flexibility? The access/privacy tradeoff

The use of secure facilities to access microdata is well established, however there is growing demand for greater flexibility in how these may be accessed. Michael Haan noted that Canada’s Research Data Centre network hosts something like 500 microdata files and that there is constant discussion about how these can be more widely used without compromising privacy. Experiments during pandemic restrictions with remote access to centrally hosted datasets in Canada and Ireland point to different options. The thin data client model being adopted by StatCan may enable more remote access with greater processing power.

Disclosure control techniques

Callum Foster noted that a variety of techniques are being used to ensure privacy in disclosure control. Microdata anonymization is a long-standing practice, as is the suppression of counts

that fall below established thresholds. More recently adopted, however, are the swapping of individual records between households, and the application of “data perturbation” to microdata files.

Lessons

8. Social norms and legal requirements regarding data privacy are in flux, largely in the direction of stronger protection.
9. It is unclear what long-term impact this will have on research use of census and other data.
10. Different risk-reward trade-offs regarding the tension between access and privacy are possible.

Appendix A: Presenters and Projects

The following information was provided by the presenters in advance of the workshop.

Session 1 – Challenges and Innovations

Presenter

Jonathan Schroeder

Research Scientist

IPUMS

University of Minnesota

Email: jps@umn.edu

Bio: Jonathan Schroeder is a geographer at the IPUMS Center for Data Integration at the University of Minnesota. He principally serves as a project manager and senior researcher for the IPUMS National Historical Geographic Information System (NHGIS), leading the integration of new U.S. census summary data and the development of time series, GIS files, and other resources. He also contributes to the IPUMS USA project by leading the development of integrated geographic variables for U.S. census microdata.

Project description: IPUMS provides census and survey data from around the world integrated across time and space. IPUMS consists of several projects and websites that each provide different types of data. Jonathan Schroeder manages one of these projects, the IPUMS National Historical Geographic Information System (NHGIS), which provides access to summary tables and time series of U.S. population, housing, agriculture, and economic data, along with GIS-compatible boundary files, for years from 1790 through the present and for all levels of U.S. census geography.

Project origin: As an academic research project

Project purpose: Make data publicly available for general use

Funders: Principal funding is "Grants from governments, including academic research support". Our institute has received some funding from each "other" listed type. We don't require any payment from users but accept donations.

Terms of funding: Subject to regular reauthorization

Concerns regarding funding: Yes. The bulk of our funding is from time-limited grants, for which we regularly need to re-apply. We have saved enough from indirect grant benefits to cover occasional unfunded intervals, but we can't do that indefinitely.

Project audience: The general public, Governments, Academic researchers, Non-profit organizations and charities, Businesses. Academic researchers and students are the principal audience of interest for our funding organizations, so we prioritize them in our work, but we welcome and support users of all kinds.

Project objectives: Sustained access to the data and materials for your audience, Linking data across time (i.e., multiple census cycles), Making data consistent over time (i.e., multiple census cycles), Realizing innovation in data dissemination (e.g., data access, usability, linking), Open access. For our international data projects, we also prioritize linking & making data consistent across countries.

Co-presenter

Humphrey Southall

Professor of Historical Geography
School of the Environment, Geography and Geoscience
University of Portsmouth
Email: humphrey.southall@port.ac.uk

Bio: Humphrey Southall is Professor of Historical Geography at the University of Portsmouth, UK. He has undergraduate and postgraduate degrees from the University of Cambridge, and also studied at Harvard University. He has directed the work of the Great Britain Historical GIS Project (GIS = Geographical Information System) since 1994, and has been a member of the Office of National Statistics' Census Advisory Group since 1999.

Co-presenter

Paula Aucott

Senior Research Associate
University of Portsmouth
Email: paula.aucott@port.ac.uk

Bio: Paula is a Senior Research Associate in the School of the Environment, Geography and Geosciences at the University of Portsmouth. She has worked on the Great Britain Historical GIS project since 2000 with a particular focus on content. Her research interests include historical GIS, gazetteers and historical land use mapping.

Project description: (Provided by Humphrey Southall.) Since 1994 I have been directing the Great Britain Historical GIS Project, a self-defined entity funded through many separate grants from the government research councils and research charities, each with its own project title. Until 2001, the main activity was mapping the changing statistical reporting units of Britain since 1801, including the c. 16,000 parishes, and computerising or otherwise obtaining digital versions of associated historical statistics. The Census of Population is our most important statistical source, but we also cover vital statistics, unemployment and the Poor Law, farm statistics and so on. In 2001 we obtained a large new grant from the UK National Lottery, enabling us to create the web site "A Vision of Britain through Time" (www.visionofbritain.org.uk) as a resource for the general public. Rather than being another download system, VoB's unique data architecture focuses on bringing together data for the user's selected locality from many and diverse geographical surveys of Britain. In particular, census data is presented primarily as local time series, both for original historical reporting units and, through GIS-based redistricting ("vector overlay"), for modern Local Authority Districts. The system was substantially extended in 2009. I am now semi-retired but we still have one full-time staff member, funded mainly from data licensing income.

Project origin: As an academic research project

Project purpose: Make data publicly available for general use

Rules regarding dissemination: No. We work only with aggregate data and associated boundary data.

Funders: Original funding from government, charities & national lottery. Now mainly data licensing & ads on website.

Terms of funding: As noted above, we now depend on licensing income from boundary data, Google ads, and some consultancy work.

Concerns regarding funding: Yes! Some of our licensing income is long-run, but not enough so we depend on one-off sales.

Project audience: The general public, Governments, Academic researchers. Our main activity is running an open access web site, and there are certainly some instances of business and third sector use. We have recently supplied digital boundary data for historical census units to the Office of National Statistics census division, and redistricted historical census data to a cross-department team working on the "Levelling-up" White Paper, so "Government" is clearly included.

Project objectives: Sustained access to the data and materials for your audience, Linking data across time (i.e., multiple census cycles), Making data consistent over time (i.e., multiple census cycles), Partial cost recovery through revenue generation

Co-presenter

Glen Hohlmann

Assistant director

Census Operations Division

Statistics Canada

Email: glen.hohlmann@statcan.gc.ca

Bio: Glen Hohlmann is the Assistant Director in the Census Operations Division at Statistics Canada responsible for the Census Dissemination Sub-Project. He holds a Master's Degree in Geography from Carleton University. Throughout his 20 year career at Statistics Canada Mr. Hohlmann has held a number of positions in the Census and Geography divisions. The 2021 Census marks his second time leading the dissemination sub-project for the Census. He is currently focused on releasing data from the 2021 Census and ensuring that census products are timely, relevant and accessible in order to better serve Canadians.

Co-presenter

David Price

Manager

Census Tabulations

Statistics Canada

Email: david.price2@statcan.gc.ca

Project description: The overall objective of the 2021 Census of Population is to ensure that Canadians have access to timely, relevant and quality statistical information on Canada's changing economy and society for informed debate, research and decision-making on social and economic issues. Through the Census of Population, Statistics Canada provides information on changes in the Canadian population and its demographic characteristics, information that serves as a basis for public and private decision-making, and research and analysis in areas of concern to Canadians. The objective of the dissemination sub-project is producing and getting relevant data into the hands of users in a timely fashion and ensuring they have the tools to use and understand them. In keeping with overall census program objectives, the census dissemination sub-project aims to promote the use of data as much as possible, making it accessible and relevant to the information needs of a broad range of data users.

Project origin: As a government initiative

Project purpose: Make data publicly available for general use

Rules regarding dissemination: No, they are all treated the same.

Funders: Government program funding

Terms of funding: Subject to regular reauthorization

Project audience: The general public, Governments, Academic researchers, Non-profit organizations and charities, Businesses, All user groups.

Project objectives: Providing access to census data to all Canadians.

Presenter

Elaine Castonguay

Assistant Director

Statistical Geomatics Centre

Statistics Canada

Email: elaine.castonguay@statcan.gc.ca

Bio: I am the Assistant Director of the Statistical Geomatics Centre at Statistics Canada. I am responsible for the Address and Geographic Register (AGR) sub-project and the Census geographic data dissemination (Geo Dissemination) task. Combined, the AGR and Geo Dissemination sub-projects support all census activities by providing the address databases and geographic areas and road networks required to support all data collection, processing, certification, weighting, evaluation and dissemination activities

Project description: This will be reflected in the workshop materials

Project origin: As a government initiative

Project purpose: Make data publicly available for general use

Rules regarding dissemination: As I am responsible for geographic data - spatial and non-spatial, there are no privacy rules, no matter who uses it.

Funders: Government program funding

Terms of funding: Permanent, ongoing funding

Concerns regarding funding: No, the development of geographic concepts for the census, the maintenance of these concepts and dissemination are required for the dissemination of demographic data and census statistics.

Project audience: The general public, Governments, Academic researchers, Non-profit organizations and charities, Businesses

Project objectives: Open access

Session 2 – Disseminating Microdata: International Perspectives

Presenter

Deirdre Lynch

Statistician

Census Dissemination

Central Statistics Office

Email: deirdre.lynch@cso.ie

Bio: My name is Deirdre Lynch and I am a statistician in the Central Statistics Office of Ireland for the previous 4 years. My background is in Economics and Geography, and I thoroughly enjoy demography. I enjoy working on Census data as it always interests me in the questions it can answer. I work on a great team of 7 people in the Census Dissemination section and we are consistently answering the public's queries for Census data. We are undertaking our next Census on April 3rd of this year, so we look forward to this period. My work mainly consists of tabulations, Parliamentary Questions, RMF creation/output approvals, planning for the next census, visualisation of data and much more.

Project description: An Irish Perspective: Census Research Microdata Files. I will present a general overview of how Ireland has and will continue to create RMFs files specific to the Census. In this presentation I will give a broad overview of the access controls, the application portal, through to how and the rule based approach used for generating the RMF and output approvals. I will then highlight where further information can be found on the CSO website.

Project origin: As an academic research project

Project purpose: Use data to solve a particular research question or policy problem

Rules regarding dissemination: Yes, it is only academic researchers and non-profit organisations (or companies procured to undertake this research) that can gain access to our Census RMF.

Funders: There is no cost associated with gaining access to our Census RMF, and it is required as part of the Census Dissemination Statisticians role to create. Therefore it is Government funds, but there is no grant available. There were however large costs with setting up the new Research Co-ordination Unit to host and control the access/applications etc. and this again was Government funding.

Terms of funding: It's a permanent role as it is included in the Central Statistics Office and is ongoing.

Concerns regarding funding: I feel this question is not applicable here given it is carried out by the Statistician in the CSO which is a permanent Civil Service occupation.

Project audience: Academic researchers

Project objectives: Preservation of original formats, Making data consistent over time (i.e., multiple census cycles), Regular reporting

Presenter

Callum Foster

Business Lead, Outputs and Dissemination

2021 Census

Office for National Statistics

Email: callum.foster@ons.gov.uk

Bio: Involved in the last 3 censuses for England & Wales, each time working on the Outputs and Dissemination part of the operation. Between 2002 and 2012, was also responsible for managing a 'Neighbourhood Statistics' website and service that pulled small area data from across government into a single service that was particularly focused on understanding areas of social deprivation. Currently working on preparations for the release of 2021 Census Outputs including some innovative ways of providing more flexibility to users.

Project description: Business Lead at Office for National Statistics, responsible for Census Outputs and Dissemination. This covers the planning and preparation of Outputs and Dissemination systems for the 2021 Census, as well as ongoing provision of support to use of data from 2011 census and earlier. We are also currently carrying out work to digitise much of our historical content with the intention of improving access to census data over time.

Project origin: As a government initiative

Project purpose: To make data publicly available for general use, and to also make other sets of data available to specific audiences in more secure and managed settings. Also to design and develop outputs and systems for the current census, alongside providing customer support, and ongoing access to data from previous censuses.

Rules regarding dissemination: Yes, in that we enable certain user audiences to have greater access to data based on certain terms and conditions.

Funders: Government program funding

Terms of funding: Subject to regular reauthorization

Concerns regarding funding: Not really, but would always like to be able to have more funding in order to be able to do more.

Project audience: The general public, Governments, Academic researchers, Non-profit organizations and charities, Businesses, Schoolchildren, specific community groups and others

Project objectives: Sustained access to the data and materials for your audience, Preservation of original formats, Linking data across time (i.e., multiple census cycles), Making data consistent over time (i.e., multiple census cycles), Realizing innovation in data dissemination (e.g, data access, usability, linking)?, Open access, Regular reporting, Involvement of specified partners

Presenter

Lara Cleveland

Project Manager, IPUMS International

Senior Research Scientist, ISRDI

Email: cleveland@umn.edu

Bio: I direct the IPUMS' International census and survey data project. In addition to overseeing data integration and metadata development for the project, I am also responsible for strategic planning and partner relations. I started working for IPUMS (Integrated Public Use Microdata Series) in 2005, coding and harmonizing historical occupational variables while finishing my graduate work. I transitioned to IPUMS International in 2008 and hold a PhD in Sociology from the University of Minnesota.

My research interests include data and methods; geographic variable harmonization and development; occupations and work; and global standardization practices. I have worked on the teams that spearheaded work on variance estimation, sample documentation, geographic harmonization, household asset indices, the secure access research data enclave, and the development of teaching materials for IPUMS.

Session 3 – Disseminating Microdata: Canadian Perspectives

Presenter

Alexandra Clemence

Chief, Orientation and Genealogy

Library and Archives Canada

Email: alexandra.clemence@bac-lac.gc.ca

Bio: Currently Chief of Orientation and Genealogy at Library and Archives Canada, I previously worked as a reference librarian in government and in academic institutions.

Project description: Library and Archives Canada holds an extensive collection of Canadian census records from 1640 to 1926, and for Newfoundland from 1671 to 1945. The Genealogy team at LAC helps contextualize and disseminate census data to the public, and plans for future releases of records.

Project origin: As a government initiative

Project purpose: Make data publicly available for general use

Rules regarding dissemination: For the most part, all users have the same access according to privacy legislation, although departmental researchers may be granted access to records containing personal information for specific tasks or projects.

Funders: Government program funding

Terms of funding: Permanent, ongoing funding

Concerns regarding funding: As census data dissemination is mandated, ongoing funding is not a concern, although the degree of indexing and the means of dissemination can be affected.

Project audience: The general public, Academic researchers

Project objectives: Sustained access to the data and materials for your audience, Preservation of original formats

Co-presenter

Sandra Sawchuk

Librarian

Mount Saint Vincent University

Email: sandra.sawchuk@msvu.ca

Bio: Sandra Sawchuk is the data services and user experience librarian at Mount Saint Vincent University Library and Archives in Halifax, NS. She currently serves as the Atlantic Regional Training Coordinator for the Data Liberation Initiative (DLI). She has an academic background in the digital humanities, and her research interests include data rescue and reuse. She recently co-authored a paper on computational reproducibility, and she is currently participating in a two-year SSHRC Partnership grant to improve access to Canada's historic census data.

Co-presenter

Susan Mowers

Research Librarian (Data)

Library

University of Ottawa

Email: smowers@uottawa.ca

Bio: Susan Mowers is a Research Data Librarian at the University of Ottawa. Through her data service and teaching activities, Susan promotes data literacy, discovery, citation, access and use. Susan has co-written a guide on citing Canadian public microdata, chaired the OCUL data community, and is continuing previous historical census project work as a participant in a two-year SSHRC Partnership grant (CCDDP).

Project description: The presentation will introduce the range and means of access to CCRI's valuable data and materials. It is hoped that the Canadian Data Discovery Data Discovery Partnership (CDDDP) project, by covering CCRI data and materials, will improve awareness and findability of this information. The CCRI (Canadian Century Research Infrastructure) project was a pan-Canadian, multi-disciplinary and multi-institutional effort to develop a set of interrelated databases centered on data from the 1911, 1921, 1931, 1941 and 1951 Canadian censuses.

Project origin: As an academic research project

Project purpose: Make data publicly available for general use (CDDPP),

Rules regarding dissemination: The CDDPP's ability is limited to make parts of historical Canadian Census data discoverable is limited according to the Statistics Act's 92-year

non-disclosure period, However, owing to the value of the CCRI historical census data and other materials, we want to make as much of the CCRI data discoverable as possible.

Funders: Government program funding

Terms of funding: One-time grant

Concerns regarding funding: Keeping the discovery portal updated and maintained.

Project audience: The general public, Academic researchers

Project objectives: Sustained access to the data and materials for your audience, Making data consistent over time (i.e., multiple census cycles), Open access

Presenter

Kris Inwood

Professor

Economics and History

University of Guelph

Email: kinwood@uoguelph.ca

Bio: Kris Inwood has been a faculty member at the University of Guelph since 1986. He is jointly appointed to the Department of History and Department of Economics. His current research examines social mobility and ethnic differences in the standard of living and incarceration using historical administrative and census data.

Project description: The Canadian Peoples/les populations canadiennes prepares Canadian complete count census data 1852-1921 for research use.

Project origin: As an academic research project

Project purpose: Use data to solve a particular research question or policy problem

Rules regarding dissemination: yes

Funders: Grants from governments, including academic research support

Terms of funding: CFI

Concerns regarding funding: Maintenance and updating of the data is a concern

Project audience: Academic researchers

Project objectives: Sustained access to the data and materials for your audience, Linking data across time (i.e., multiple census cycles), Making data consistent over time (i.e., multiple census cycles), Realizing innovation in data dissemination (e.g, data access, usability, linking)?

Presenter

Michael Haan

Associate Professor

Sociology

Western University

Email: mhaan2@uwo.ca

Bio: Dr. Michael Haan (PhD, University of Toronto, 2006) is an Associate Professor and RDC Academic Director at Western University. From 2005-2010, Dr. Haan held an academic appointment at the University of Alberta, where he was Winspear-Archer Research Fellow in Immigration and Social Policy. From 2010-2015, he was Canada Research Chair in Population and Social Policy at the University of New Brunswick. He is also research associate at the Prentice Institute for Global Population and Labour at the University of

Lethbridge, and at the McGill Centre for Population Dynamics. His research interests intersect the areas of demography, immigrant settlement, labour market integration, and data development. Dr. Haan is widely consulted by provincial and federal governments for policy advice in the areas of immigration, settlement services, the Canadian labour market, and population aging. Dr. Haan is currently investigator or co-investigator on over ten million dollars of research focused on immigrant settlement, developing welcoming communities, and identifying the factors that predict successful retention of newcomers. Since receiving his PhD in 2006, he has already published over 75 articles and reports on these topics.

Project description: Research Data Centres are secure physical environments available to accredited researchers and government employees to access anonymized microdata for research purposes. Researchers have direct access to a wide range of anonymized (de-identified) survey, administrative, and integrated data. The COVID-19 pandemic has made clear the need for Statistics Canada to transition from the RDCs' physical infrastructure to new distributed modes of access. A plan to address limitations of the current physical infrastructure is currently underway.

Project origin: As an academic research project

Project purpose: Use data to solve a particular research question or policy problem

Rules regarding dissemination: No

Funders: Government program funding

Terms of funding: Subject to regular reauthorization

Concerns regarding funding: No

Project audience: The general public, Governments, Academic researchers, Non-profit organizations and charities, Businesses

Project objectives: Sustained access to the data and materials for your audience, Preservation of original formats, Linking data across time (i.e., multiple census cycles), Making data consistent over time (i.e., multiple census cycles), Realizing innovation in data dissemination (e.g., data access, usability, linking), Open access, Partial cost recovery through revenue generation.

Appendix B: Overview of Census Discovery Platform Features as Demonstrated

This appendix summarizes the speakers' comments regarding the functionality of their Census data platforms.

For aggregate census data

National Historical Geographic Information System (NHGIS)

<https://www.nhgis.org/>

Jonathan Schroeder

Targeted at users who want to download data.

Easy data retrieval steps:

- 1 Landing page is literally a Get data button that takes users to 4 filters: Years, Geographic levels, Datasets and Topic.
- 2 Immediately after this page, users are taken to dataset results by Source data, Time series tables, and GIS files.
- 3 Time series tables are built on harmonized census variables and value codes, and geography. These table lists include metadata on the type of geographic integration, years, geographic levels, short descriptive table names.

In development,

- API support, R support
- Looking at new tabulations from historical microdata

Email: gbhgis@port.ac.uk

Mention of new interface of the **US Census Bureau**

<https://data.census.gov/cedsci> - Targeted at all users. Some difficulties, e.g., finding basic tables.

Vision of Britain

<https://www.visionofbritain.org.uk/>

Humphrey Southall

- Targeted at the general public, and not researchers or others wanting to map data
- Designed to find basic information immediately

Access:

1-a - *Main means of access: their hierarchical geographic structure is exposed to Google. Most of their users come to their site from a web search - they land on the "Place page"*

1-b - Their website landing page: search by place (via name or map).

2 Place page

- immediate place information, e.g., summarized gazetteer information
- Sidebar (the spine). Option: "Units and statistics"

- 3 Units and statistics links page. Units = administrative units. 2 tabs: Units named after this place (by themes), Units covering this place
 4. Select unit (geographic level)
 - 5 Place through time page, See tabs: Historical statistics, Key Findings ...
 - 6 Key findings page, Views: Chart, Table, Definitions and source information.
- Project funding supported focus group testing of a demo site. Major funding was received in 2001 to build the website.
 - Delivers very popular local area census time series. Users prefer these to cross section data. Site complexity is hidden from users.
 - Geographic structure is exposed to google, landing users on its Place pages. About 300 time series based on selected published tables
 - All data is in one database @19,498,650 rows (statistical units). They took complex administrative geographic units and grouped them into "places" primarily based on name.

Census of Population (Statistics Canada)

<https://www12.statcan.gc.ca/census-recensement/index-eng.cfm>

Glen Hohlmann, David Price

- Targeted at all users
 - Includes, but not limited to,
 - [Census Program Data Viewer](#) (interactive map, 2016, 2021)
 - by year: Profiles & Data tables
- (A) Formats:
- 1) Beyond 20/20. Why? (a) Usability: Ease of use / Supports data use for a wide audience, (b) Feedback received from Census 2021 consultations to not get rid of B2020 e.g., for the Data Files (i.e., multidimensional tables), (c) The 2021 Census Data tables will contain many billions of cells this year
 - 2) CSV and SDMX (all tables). Why? a) Neutral format / versatile (CSV), b) Preservation formats
- (B) Usage. The Census Profiles and Data Tables are well-used. The Census Profiles receive approximately 5 million views a year.

Geography (Statistics Canada)

<https://www.statcan.gc.ca/en/geography>

Elaine Castonguay

- (A) [Geography page](#)
- Includes intercensal products
- (B) [Census Geography page \(2021 Census\)](#)
- Spatial information (GIS) products
 - Maps
 - Attribute information products
 - Reference documents & more
- (C) [Standard geographic classification \(SGC\)](#)
- Standard geography, harmonized over time, includes Reference Maps. The SGC (2021) is in its 11th edition.
 - The main goal is to support the Census

For census microdata - International (includes mention of aggregate census data)

Central Statistics Office (Ireland)

<https://data.cso.ie/#>

Deirdre Lynch

CSO open data

A) PxStat (aggregate census data tables)

- Targeted at all users from [CSO web site](#)

1. Browse
2. People and Society
3. various Censuses of Ireland

Notes:

- CSO has moved away from 2 and 3 dimensional tables with geography
- Every year about 1,200 tables
- [PxStat](#) is a Data Dissemination Management System for publishing Statistics in Open Data formats

B) Special tabulations by small geographies created for public organizations

1. [Census 2016 Small Area Population Statistics](#)
2. [SAPMAP](#)

- Targeted at power users and users interested in their area

- Popular map interface

C) Anonymized Microdata Files (AMF) via [Irish Social Science Data Archive](#) and via [IPUMS-International](#)

Office for National Statistics (UK)

[Census 2021](#)

Callum Foster

Demonstrated two innovative tools for providing flexible aggregate data tables from underlying census microdata. The tools address the needs of two different user groups, and make more data available than the standard, or ready-made outputs (which however, still provide blanket geographic coverage (all areas), are good for their level of comparability and for users who may not want to build their own tables.

INNOVATIVE TOOLS:

A) First tool: flexible tables

- Targeted at a more expert statistician user and uses Contabula to which the ONS contributed
1. Choose your population (e.g., "People with households, England and Wales")
 2. Choose your geographic level, e.g., "Lower layer super output area". When you click on the area, "Your table" displays the number of cells / areas, and the name of the dataset (People with households, England and Wales)
 3. Select specific (by name) areas, all areas (whole population), or all areas within a larger area.

4. Choose your variables (and breakdown detail: number of categories)
 5. (Disclosure control process starts, and *interface displays* "Your table" details including the disclosure success rate as the number of areas that pass out of total number of areas, and an infographic)
 6. In relation to your success rate, you can adjust the level of chosen detail and add or take away variables.
 7. You can view or download your data
 8. Several download options include, (a) changing your dataset, (b) geographic level (e.g., Middle Layer Super Output Level or Local Authority), (c) geographic area, and (d) variables. Seamlessly calling up the relevant selection screen with your success rate information. Filter table.
 9. Supports different output formats
- (B) Second tool is a potential interface that sits on top of the first tool.
- Population Explorer is an exploratory tool
 - As an exploratory tool it brings data to life, and quickly provides insights from the data in a more accessible form
 - It is targeted at a more policy influencer type user who needs data to support their everyday work, but doesn't use data everyday.
1. Select a topic, e.g., Arrival in UK, length of time, and passports held
 2. Shows percent of total population in England and Wales, median age, age profile, a choropleth map, and top 5 areas with highest and lowest percentages
 3. Shows key indicators for General health, Marital status, Social grade, Economic activity, Distance to work, Mode of travel to work
 4. Can scroll over the map to identify geographical differences (% of population)
- (C) Third tool - Flex-It
- Targeted at the public
 - ONS to introduce Flex-It tools for various ready-made standard Census 2021 tables so that, e.g., users can swap variables.

IPUMS-International

<https://international.ipums.org/international/>

Lara Cleveland

- Narrower audience than other IPUMS projects: Not to give data to for profit entities
- Database user tools and standards
1. Access via ...
 - Different modes of dissemination for academic users
 - gives microdata because academic users want the microdata on their computers
 - Metadata is essential for responsible use (rely on good sample design documentation)
 - Increasing interest shown by the policy sector
 2. Facilitating and managing open dissemination via ...
 - Harmonization / standardization (e.g., variables) is extremely helpful and efficient
 - Use registration to alert users to changes/errors, make connections
 - Web automation for fulfillment of requests

- Standardized linking fields - especially geography
- Three codebook options:
 - General level of codes (IPUMs harmonized codes),
 - Detailed level of codes (IPUMS harmonized codes + the comparable country codes),
 - Fully harmonized codes (only the cross-censuses *common* harmonized codes)

For census microdata - Canada

Library and Archives Canada

17 Census databases, 1825 - 1926

<https://www.bac-lac.gc.ca/eng/census/Pages/census.aspx#b>

Alexandra Clemence

Census metadata includes:

- column headings
- instructions to enumerators and other resources
- section about relevant issues & problems..

Database searching

- Each database currently has its own search tool. Typical search tool:

1. Search screen

- Searching by: keywords, names, ages, provinces, names and numeric identifiers of Districts and Sub-Districts, page numbers.
- Includes search help

2. Click on the entry

3. Digitized image appears (click on the image to see the full image for the census return)

4. Click the item number for archival reference and to suggest a correction

CCRI (Canadian Century Research Infrastructure)

Covering various web sites

Susan Mowers and Sandra Sawchuk

Demonstration of [CCRI Guide \(preliminary version\)](#) covering CCRI resources found on various platforms and where to find them.

1. The five CCRI Canadian Census microdata files ("databases"), (1911, 1921, 1931, 1941, 1951) form the core of CCRI's data
2. The five microdata files feature harmonized census variables and harmonized geography.
3. The guide also covers metadata, spatial data, reference maps, and introductory selected published Census tables.
4. The CCRI archive "web gateways" include:
 - [University of Alberta](#)
 - [University of Alberta Dataverse](#)
 - [l'Université du Québec à Trois Rivières](#)

- [University of Toronto](#)
 - [University of Ottawa](#)
 - Alternate platforms: data repositories housing CCRI resources
5. Further CCRI partner university websites include
- [The lead project website - University of Ottawa](#)
*Note, the above website was unavailable at the time of writing.
Please note instead the [2019 website capture](#) available via the
Internet Archive Wayback Machine.*
 - [CCRI York University Centre website](#)
 - [CCRI-ICRS 1911 Sample of the 1911 Canadian Census website
\(Université du Québec à Trois-Rivières\)](#)