# DATA ACCESS DIVISION NEWSLETTER

## WINTER 2020 EDITION

### Greetings and Happy New Year!

The Data Access Division would like to wish you a Happy New Year! We want to thank you and extend our warmest wishes for the season to all our clients and friends who have made our data access progress possible. Our thoughts continue to be with all those who have been affected by the COVID-19 pandemic. We are pleased to see that our community is staying strong and optimistic through these difficult times.

We hope you enjoyed the holiday season with your loved ones, and we look forward to the continued support in the coming year.

### DATA ACCESS DIVISION:
### NEW DIVISION NAME

The Microdata Access Division is excited to announce that, with its continued rebranding and reorganization, it has changed its name to the Data Access Division (DAD). The change went into effect in November 2020 to reflect structural changes happening within the Division.

In the new year, DAD will continue to provide increased and timely data access in new and innovative ways. The Division will also keep offering high-quality services for researchers and clients to best serve their data needs across the country.

For more information, please visit the Data Access Division website.

# DATA ACCESS DIVISION NEWSLETTER

## SELF-SERVE ACCESS

### DATA LIBERATION INITIATIVE TEAM UPDATES

The Data Liberation Initiative (DLI) held its first virtual training sessions from November 23 to 27. These sessions were hosted by the Professional Development Committee (PDC) and Statistics Canada. In all, 155 participants were registered, with an overall attendance of 73.4%. A wide variety of topics were covered; the schedule can be found on the DLI training website. If you could not attend, the presentations and training resources can be found in the DLI Training Repository. We would like to thank all the presenters, students, translators, note takers and moderators, as this would not have been possible without them volunteering their time and expertise.

### WHAT STUDENTS HAD TO SAY

*"Hearing about future StatCan projects was valuable. I'm pleased to see how concepts will be updated for the next census (especially around sex and gender). Improving the census is an ongoing process, and while I don't think there were any surprises among the changes discussed, I'm starting to get excited for 2021. I am also thrilled to hear about the ways that StatCan is making microdata secure and more readily available. I'm a student librarian who occasionally performs reference duties, and even I will experience the tangible value of this service—especially when business data are implemented!"* - **Dan Philips**

See the full comment here.

An article written by Barbara Znamirowski, "DLI National Training: Accessing data and building a community," provides a great overview of the week-long training.

### PROFESSIONAL DEVELOPMENT COMMITTEE

The PDC will send a call-out to the listserv in the new year for two volunteers to represent the Quebec and Ontario regions. The DLI team and the PDC would like to take this opportunity to welcome new members to the PDC and thank those who are stepping down.

**New members:**

- Atlantic Region – Margaret Vail
  St. Francis Xavier University
- Quebec Region – Emanuela Chiriac
  Université du Québec en Outaouais
- College Representative – Caleb Domsy
  Humber College

**Members stepping down:**

- Carolyn DeLorey
  St. Francis Xavier University
- Caroline Patenaude
  Université de Montréal
- Gaston Quirion
  Université Laval (in May 2021)
- Deena Yanofsky
  University of Toronto

The PDC has put together a working group to update the DLI Contact and Alternate's Survival Guide. The new guide will be in line with the StatsCan Web Redesign Project. As part of this initiative, a notice was sent out to the community for comments.

# DATA ACCESS DIVISION NEWSLETTER
## SELF-SERVE ACCESS

## PUBLIC USE MICRODATA FILES ONLINE PROJECT

The project to put public use microdata files (PUMFs) online in a downloadable format is underway. As part of this project, digital object identifiers (DOIs) will be assigned to PUMFs. Since May, Statistics Canada has released 20 PUMFs, including 12 COVID-19 PUMFs.

Colectica, which uses the Data Documentation Initiative standard, has been approved by the agency as the data documentation tool. Once it is implemented and becomes widely used across the agency, it will be easier to create PUMFs. This should increase the number of PUMFs released every year.

## REAL TIME REMOTE ACCESS INITIATIVES:
## EXTENDING ACCESS

Research data centre (RDC) researchers have had their access to the Real Time Remote Access (RTRA) program extended to March 31, 2021. The RTRA team has compiled the following metrics and usage report:

- 48 users submitting
- 1,455 submissions
- 922 successful submissions
- 2,779 tables output

A list of all DLI products is available on the website:

www.statcan.gc.ca/eng/dli/dli-products.

### DATA RELEASES TO DLI SINCE OCTOBER 2020:

- Labour Force Survey files—monthly
- Employment Insurance Coverage Survey - 2019
- Canadian Student Tobacco, Alcohol and Drugs Survey – 2014, 2016 & 2018
- Social Policy Simulation Database and Model (SPSD/M), COVID-19 glass box, version 3.0
- Crowdsourcing: Impacts of COVID-19 on Canadians Experience of Discrimination – 2020
- Canadian Perspectives Survey Series 4: Information Sources Consulted During the Pandemic Public Use Microdata File
- TLAC Standard Tables 2020-2021
- Interprovincial and International Trade Flow 2017
- Supply and Use Tables 2017

A list of all RTRA products is available on the website:

https://www.statcan.gc.ca/eng/rtra/surv.

### DATA RELEASES TO RTRA SINCE OCTOBER 2020:

- Registered Apprenticeship Information System (RAIS)
- Labour Force Survey—released monthly

# DATA ACCESS DIVISION NEWSLETTER
## RESEARCH DATA CENTRES

### RESEARCH DATA CENTRE UPDATES

RDCs are once again operational after the shutdown at the beginning of the pandemic. Only two centres have yet to reopen, and the remainder are operating at reduced capacity with COVID-19 safety measures in place. As the pandemic resurgence is underway, staff are preparing in the event a second round of closures is required in some jurisdictions.

Work is ongoing to help ensure that research with Statistics Canada data can continue if closures occur. For example, DAD is preparing to pilot a new virtual data access platform. In this project, Statistics Canada microdata will be made accessible securely outside RDC facilities using a Statistics Canada cloud infrastructure platform. This project will involve four universities, to start, but will be implemented with broader contingency planning in mind if RDCs in other universities have to close because of the pandemic. Approximately eight RDC research teams accessing eligible survey data will be offered this new opportunity in January 2021. The lessons learned from this project will help the program determine next steps for expanding access more broadly by the end of 2021.

Work continues on the project to modernize the RDC information technology (IT) infrastructure by updating the network architecture and moving to a centralized virtual environment. This initiative, referred to as the vRDC, will greatly improve computing resources for academic researchers and improve the reliability of the system. Remote access to the vRDC will follow the same access framework that is used in the Virtual DataLab (vDL) environment. The rollout is anticipated to start in late 2021.

In spring 2020, a high-level task force co-chaired by Martin Taylor (Director, Canadian Research Data Centre Network) and Jacques Fauteux (Assistant Chief Statistician, StatCan) was created to develop short- and long-term data access strategies for academic researchers. Its mandate is to align different initiatives, including the vRDC and vDL.

In October, the Canadian Research Data Centre Network, in partnership with Statistics Canada, held a very successful 20th anniversary virtual conference. Attendance was at an all-time high for the conference, with more than 950 participants registering. Feedback was extremely positive.

### NEW RESEARCH DATA CENTRE HOLDINGS

A total of 31 products were added to microdata holdings in the third quarter of 2020. These include two new surveys (Crowdsourcing: Impacts of COVID-19 on Canadians' Experiences of Discrimination, and the 2019 Canadian Health Survey on Seniors), as well as updated survey cycles and linked and administrative datasets.

### PARTIAL LIST OF NEW DATA RELEASES FROM SEPTEMBER TO DECEMBER 2020

- Labour Force Survey (LFS) COVID Fast Track Option (FTO) 2020 - September
- Canadian Perspectives Survey Series 3 – 5 (CPSS) 2020
- Survey on Individual Safety in the Postsecondary Student Population (SISPSP) - 2019
- Canadian Housing Survey (CHS) 2018 linked to Tax data, Social Inclusion, Proximity Measure, Income Dispersion and Historical addresses
- Longitudinal Administrative Databank (LAD) data linked to Discharge Abstract Database (DAD) 1997-2016

- Canada Education Savings Program (CESP) linked to T1 Family File (T1FF), Census 2016 and LAD
- Canada Mortgage and Housing Corporation (CMHC) Custom Data Files linked to 2017 Canadian Survey on Disability
- Canadian Cancer Registry (CCR) 2017
  - The International Agency for Research on Cancer (IARC)
  - Surveillance Epidemiology and End Results (SEER)
- Provisional death data January to September 2020 (monthly extract from Vital Statistics dataset)

For a complete list of data available in RDCs and government access centres, visit www.statcan.gc.ca/eng/rdc/data.

## GOVERNMENT DATA ACCESS:

### FEDERAL RESEARCH DATA CENTRES

In April 2020, DAD assumed the responsibility of providing access to Statistics Canada's business data holdings to deemed employee researchers through data access programs. Throughout the fall, the Government Data Access team worked to integrate business data access under the Federal Research Data Centre (FRDC) umbrella. Currently, the FRDC operates two data centres at Tunney's Pasture under this umbrella, the Social Data Access Centre and the Business Data Access Centre (formerly known as the Canadian Centre for Data Development and Economic Research). Both will continue to operate for the foreseeable future to facilitate maximum accessibility to deemed employees, while ensuring appropriate physical distancing. In the future, the two centres will be fully integrated into one physical location to serve all federal government users.

As part of the transition, website content is being updated to ensure that the processes for requesting access to business data are reflected on the DAD website and within the Microdata Access Portal (MAP) instructions. Researchers interested in accessing business data are encouraged to consult with a Statistics Canada subject-matter expert before submitting a proposal through the MAP. Please contact the team to be connected with an expert.

### EXPANDING ACCESS TO BUSINESS DATA

DAD staff are continuing to expand access to business data outside the data centres located at Statistics Canada's head offices in Ottawa. To this end, the Bank of Canada remote site opened in September to facilitate on-site access to Bank of Canada staff. In addition, plans are being solidified to move access for academic clients to the university-based RDCs in the near future.

### PROVINCIAL SECURE ACCESS POINTS

In fall 2020, two secure access points opened in the provinces to facilitate access to Education and Labour Market Longitudinal Platform data. One is at the Office of Statistics and Information in Edmonton, Alberta, while the second is in Victoria, British Columbia, at BC Stats. In addition, an office at the Ontario Ministry of Finance in Toronto will be opening soon. Opportunities to open additional provincial sites continue to be sought to facilitate access outside Ottawa. Discussions are underway with the provinces of Manitoba and Prince Edward Island. For more information about these initiatives, please contact statcan.mad-ssap-dam-pass.statcan@canada.ca.

# DATA ACCESS DIVISION NEWSLETTER
## MODERNIZATION OF ACCESS

## PILOT PROJECTS AND TESTING

The Virtual DataLab (vDL) team is continuing to work with the Data Analytics as a Service (DAaaS) team to test platforms in a cloud environment with various tools to facilitate data staging and research. The platform evaluations and improvements are still in progress, ensuring that users will have access to the environment and tools required to complete their research. Additionally, the DAaaS team is working to incorporate various specifications for a search tool, which will help researchers identify relevant available datasets.

The vDL development and pilot schedules are regularly reviewed, in collaboration with IT teams, to ensure continued testing and troubleshooting before onboarding users into the vDL. The vDL team has identified several pilot projects using low- to medium-sensitive confidential microdata on the cloud environment. The users will be onboarded once the appropriate agreements have been signed, as per the approved governance structure.

The established pilot projects help evaluate the nuances of onboarding different types of researchers. In addition, the lessons learned from each pilot will be monitored and resolved or mitigated for the following project, and this will help guide the team towards a production environment.

Once the vDL team has successfully conducted pilot projects using data of low to medium sensitivity, in 2021, additional pilot projects may be conducted using data of medium to high sensitivity, as this is more representative of the data that researchers typically use. Statistics Canada's Strategic Management Committee approved this strategy on November 25, 2020.

## ACCESS FRAMEWORK APPROVED

In line with the goal of modernizing data access, the vDL team was granted approval to introduce a new access location: authorized workspaces. An authorized workspace can include a deemed employee's place of work, as well as a private space within their personal residence; this is timely, given the current public health situation. A summary of the privacy impact assessment for this new location of work will soon be available on the Statistics Canada website.

Along with introducing an accreditation model for trusted researchers and organizations, the VDL team developed a table for approved access locations. This table uses the sensitivity of the data, along with the researcher's and organization's accreditation scores, to determine the options for approved access locations for a given project. This table was presented to and approved by Statistics Canada's Security Coordination Committee and Data Out: Services, Access, Dissemination and Communication Committee. It will be implemented during the VDL pilot projects.

## VIRTUAL DATA LAB PROJECT UPDATES

The VDL will greatly improve access to statistical information for researchers by providing 24/7 remote access to data housed at Statistics Canada
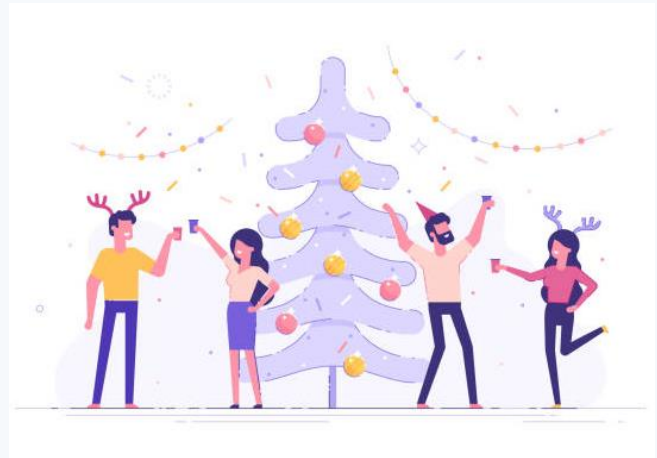
# DATA ACCESS DIVISION NEWSLETTER
## MODERNIZATION OF ACCESS

using a secure IT connection and a protected cloud environment. Progress is ongoing on a number of key initiatives to increase virtual data access and promote collaboration. These include the development of analytics platforms and the continued assessment and development of the Client Relationship Management System (CRMS) and the Microdata Search Tool.

The DAD CRMS pilot project is finished, and the CRMS corporate project is now with the Dissemination Division. Under Mathieu Laporte's guidance, Dissemination will explore the capabilities and functionality of the internal portal for CRMS.

A new production schedule is currently being developed to integrate DAD requirements. A Microsoft Dynamics consultant is expected to begin working in collaboration with Dissemination in the new year. Stay tuned!



**Questions or comments? Visit our website:** www.statcan.gc.ca/eng/help/microdata.

**Check out the StatCan Blog:** www.statcan.gc.ca/eng/blog.

**Don't forget to follow us on social media!**

Facebook | Instagram | LinkedIn | Reddit | Twitter | YouTube

Statistics Canada / Statistique Canada

Canada