

Open Access and Open Data: Challenges and Opportunities in Research

Open Day @ McMaster

Oct 25, 2018

Vivek Jadon, Mills Library

Laura Banfield, Health Sciences Library

Overview

- Defining Open Data
- Open Data in Canada
- Opportunities
- Strategies for Identifying Open Data
- Challenges
- Open Data Tools and Services
- Support the Open Data Community

Defining Open Data

- *“Open data is data that can be **freely used, re-used and redistributed** by anyone - subject only, at most, to the requirement to **attribute and share-alike.**”*
(<http://opendatahandbook.org/guide/en/what-is-open-data/>)
- *“Open data is data that anyone can **access, use or share**”* (<https://theodi.org/what-is-open-data>)
- Microdata vs Statistics: Is Open Data a Microdata or Statistics or both?
- Open data may include non-textual material such as maps, genomes, connectomes, chemical compounds, mathematical and scientific formulae, medical data and practice, bioscience and biodiversity etc.

Open Data Features

Complete Primary Timely Accessible

Machine-readable

Non-proprietary License-free

Reviewable Discoverable

Permanent Access Redistribution Reuse Description

Metadata Attribution Integrity

Absence of Technological Restriction

Open Data in Canada

- Open data in Canada dates back to 1970s with the sharing of Satellite imagery data, Data Liberation Initiative in 1990s, GeoGratis and Geobase in early 2000s.
- **Federal**
 - Design and implementation of “data.gc.ca” portal in 2012 as part of Canada’s action plan on Open Government, as well as Geospatial data made freely available through GeoGratis (NRCAN)
 - In 2014, government announced 2nd Action plan on Open Government, including the following open data commitments:
 - [Open Data Canada](#)
 - [Canadian Open Data Exchange \(ODX\)](#)
 - [Open Data for Development \(OD4D\)](#)
 - [Open Data Core Commitment](#)

Open Data in Canada

- **Open Provinces:**

- Alberta: [Open Government Portal](#)
- British Columbia: [DataBC](#)
- Ontario: [Ontario open data](#)
- Quebec: [Open Data: Government of Québec](#) (French only)
- Nova Scotia: [Open Data Portal](#)
- Saskatchewan: [Open Data Saskatchewan](#)
- Prince Edward Island: [Government of Prince Edward Island - Search open data](#)
- Newfoundland & Labrador: [Open Government](#)
- Northwest Territories: [OpenNWT](#)

Open Data in Canada

- **Open Municipalities**

- City of Hamilton: [Open and Accessible Data](#)
- City of Burlington: [Open Data Burlington](#)
- City of Niagara Falls: [Open Data](#)
- City of Brampton: [OpenGov Brampton](#)
- City of Mississauga: [Mississauga Data](#)
- City of Ottawa: [Open Data Ottawa](#)
- City of London: [Open Data](#)
- City of Toronto: [Open Data - Accessing City Hall](#)
- City of Waterloo: [Open Data Waterloo](#)

Complete list of Open Municipalities can be found on [Open Government Portal](#)

Open Data in Canada

- **National Open Data Initiatives**

- [Polar Data Catalogue](#)
- [The Grand River Conservation Authority \(GRCA\)](#)
- [Nunavut Geoscience](#)
- [Northwest Territories Geoscience](#)
- [Geomatics Yukon](#)
- [GIS Data Layers: Geographic Information for PEI](#)
- [St. Lawrence Global Observatory](#)
- [Manitoba Land Initiative](#)
- [GeoNOVA Portal](#)

Open Data in Canada

- The [Open Data Barometer](#) by WVV Foundation looks at 30 governments and how they are publishing and using Open Data for accountability, innovation and social impact.
- In 2016, Canada was considered a “Leader” in Open Data Initiatives among 30 governments and was ranked 2nd in the world.
- As of July 2018, Canada was ranked 7th.

Opportunities

- Readily available
- Time and cost saving
- Support for Research and Innovation: [Canadian Open Data Experience \(CODE\)](#), applications of Open Data: [Apps Gallery](#)
- Accountability and Community Engagement
- Education
- Governance

Strategies for Identifying Open Data

1. Contacting Campus Data Experts/Services
2. Conducting a Literature Search
3. Working with Registries, Repositories, and Portals
4. Using Internet Search Engines

Strategies for Identifying Open Data

• 1. Campus and Research Community Data Experts/Services

- Library Data Services, Mills Library
- Institutional Research and Analysis (IRA)
Data Services

The **Library Data Service** promotes and facilitates the use of specialized electronic research data products for McMaster affiliated users.

Overview

Library Data Service serves as a central point of access to a wide assortment of digital numeric research data files. The collection covers a broad range of subject areas with an emphasis on Social Sciences, Health Sciences, Business, and Economics. Access to the numeric microdata files is restricted to current McMaster University faculty, staff and students for the purposes of teaching, research and private study only.

Microdata

- Microdata is the raw data used to calculate statistics and can be manipulated with software such as SAS or SPSS
- As a member of two major consortia - Statistics Canada's *Data Liberation Initiative (DLI)* and *Inter-University Consortium for Political and Social Research (ICPSR)*, Library Data Service regularly acquires new microdata files.
- The DLI microdata files contain information on Canadian statistics collected from different surveys conducted by Statistics Canada such as *National Population Health Survey*, *National Longitudinal Survey of Children and Youth*, *Survey of Labour and Income Dynamics* etc.
- The ICPSR microdata files contain mainly US survey statistics such as the *General Social Survey*, *Current Population Survey*, *American National Election Study*, *2000: Pre- And Post-election Survey* etc.

Aggregate Data/Statistics

- Aggregate data consists of facts or figures and is used to prove a point using totals, sums or averages.
- Library Data Service provides access to aggregate data (Canadian, US & International) through the self-serve websites. Please visit our [Data/Statistics Collection](#) pages to get access.

DATA CONSULTATION REQUEST

To book an appointment with the Data Specialist (Mills Library, 1st floor, L104/C), please contact Vivek Jadon at 905.525.9140, ext. 23848 or email libdata@mcmaster.ca.

RELATED SPACES

MAPS, DATA, GIS

Library Data Service

McMaster University Office of Institutional Research and Analysis

UNDERGRADUATE PLAN MOVEMENT

- Entry-level Plan
- Other
- Science
- Engineering
- Health Sciences
- Business
- Arts and Science
- Humanities
- Social Sciences
- Medicine
- Interns and Residents

CLICK TO FILTER

BI & IRA PORTALS DATA PUBLICATIONS ACCOUNTABILITY FACT BOOK ABOUT US

Institutional Research and Analysis

Reaching out to known Groups/Labs, Research Centres, Researchers

Strategies for Identifying Open Data

- **2. Literature Search**

- Medline, Web of Science Core Collection, BIOSIS
 - Scopus and others
- Articles (may identify)
 - Supplementary Files
 - Links to Repositories
 - Named Datasets and Sources
 - Labs and Research Centres
- Reference Lists

PubMed

“Dataset [Publication Type]

- Works consisting of organized collections of data, which have been stored permanently in a formalized manner suitable for communication, interpretation, or processing.
- Year introduced: 2014”

“Datasets as Topic (MeSH)

- Subject matter related to the curation of data from research projects, stored permanently in a formalized manner suitable for communication, interpretation, or processing.
- Year introduced: 2015”

Consider using keywords for dataset*, data set*, open data*

Strategies for Identifying Open Data

3. Registries, Repositories, and Portals

Registries

- re3data.org
- Google Public Data Directory: <https://www.google.com/publicdata/directory>
- Registry of Open Data on AWS: <https://registry.opendata.aws/>

Repositories

- Scholars Portal Dataverse: <https://dataverse.scholarsportal.info/>
- GapMinder Data: <https://www.gapminder.org/data/>
- DataHub: <https://datahub.io/>

Portals

- Open Data from Government of Canada: <https://open.canada.ca/en/open-data>
- U.S. Government's Open Data: <https://www.data.gov/>
- UN Open Data portal: <http://data.un.org/>
- World Bank Open Data portal: <https://data.worldbank.org/>

Strategies for Identifying Open Data

4. Internet Search Engines

Develop Search Language

- *dataset and _____*
- *open data and _____*
- *public data and _____*
- *open datasets and _____*
- *publicly available data and _____*
- *open datasets and research*

Result Sets

- Repositories
- Listings and Directories
- Reports
 - Data Sources
- Datasets
- Centres and Labs

Google open data and asthma

All News Images Maps Videos More Settings Tools

About 10,900,000 results (0.42 seconds)

asthma data on data.world | 11 datasets available
<https://data.world/datasets/asthma>
 Find open data about asthma contributed by thousands of users and organizations across the ... California asthma emergency department visit rates by ZIP code.

asthma - Datasets - European Data Portal
<https://www.europeandataportal.eu/data/en/dataset?tags=asthma>
 This dataset contains information on Stockport residents who are smokers and/or have asthma, aggregated by LSOA. The data provided for each item is the ...

NIH Expands Open-Access Dataset of Genetic and Clinical Data to ...
<https://www.nih.gov/.../nih-expands-open-access-dataset-genetic-clinical-data-include-...>
 Dec 15, 2008 - The National Heart, Lung, and Blood Institute (NHLBI) of the National Institutes of Health has expanded its collection of genetic and clinical data first made available in 2004. The new data includes information on genetic and clinical data first made available worldwide last year, to include clinical and genetic information on ...

Open Data Portal Provides California Asthma Data
www.rampasthma.org/archives/13206
 In collaboration with the state asthma program California Breathing Program, the Open Data Portal provides datasets to the Open Data Portal. 2012 asthma emergency department visit rates by ZIP code.

Asthma ED Visit Rates (LGHC Indicator) - Data
<https://data.chhs.ca.gov/dataset/asthma-ed-visit-rates-lghc>
 This dataset contains counts and rates (per 10,000 residents) of emergency department visits for asthma. These data include emergency department visits by ZIP code. Contact Email, opendata@cph.ca.gov.

Canadian Chronic Disease Surveillance System
<https://open.canada.ca/data/en/dataset/9525c8c0-554a-4661-b000-000000000000>
 Open the TBS Open Data Portal ... hypertension, mental illness and AMI is 1154120 ... Tasha, the Open Government Team.

VCAMS Proportion of young people with current asthma
<https://www.data.vic.gov.au/data/dataset/vcams-proportion-of-young-people-with-current-asthma>
 Nov 16, 2018 - This data is collected as part of the Victorian Survey of Children and Young People (VCAMS).

Google public data and asthma

All News Maps Images Videos More Settings Tools

About 84,000,000 results (0.37 seconds)

asthma data on data.world | 11 datasets available
<https://data.world/datasets/asthma>
 Find open data about asthma contributed by thousands of users and organizations across the ... OPEN Dataset • Updated Apr 26, 2017 • Public Domain License.

CDC - Asthma - Data and Surveillance - Asthma Surveillance Data
<https://www.cdc.gov/asthma/astmadata.htm>
 Sep 8, 2016 - Asthma - Data, Statistics, and Surveillance ... TTY: 888-232-6348. Email CDC-INFO@hhs.gov. U.S. Department of Health & Human Services - HHS/Open

Asthma in Canada - Data Blog - Chronic Disease Infobase | Public Health Agency of Canada
<https://infobase.phac-aspc.gc.ca/datalab/asthma-blog-en.html>
 May 1, 2018 - Asthma is one of the most common chronic respiratory diseases in Canada. ... You're reading our Data Blog, a new way of discovering public health data.

Asthma - Data.gov
<https://catalog.data.gov/dataset?tags=asthma>
 This dataset contains case counts, rates, and confidence intervals of asthma hospitalizations in Chicago, by year, 2000-2010.

NIH Expand Asthma Data
<https://www.nih.gov/news-events/press-releases/studies-asthma>
 Dec 15, 2008 - A new study from the National Institutes of Health (NIH) shows that asthma symptoms can be improved by using a combination of a long-acting beta2-agonist (LABA) and an inhaled corticosteroid (ICS).

Asthma Metrics
<https://data.mai.gov.au/dataset/asthma-metrics>
 Age-adjusted rates of asthma hospitalizations improve public health.

Fast Facts about Asthma
<https://www.canada.ca/en/public-health/services/publications/diseases-conditions/asthma-chronic-obstructive-pulmonary-disease-canada-2018.html>
 Jul 23, 2015 - Fast Facts about Asthma.

EUROPEAN DATA PORTAL

European Data Portal > Datasets

What we do Data Providing Data Using Data Resources

Filter by location Clear

Search Location...

Search datasets...

13 datasets found Order by: Last Modified

Tags: asthma

Gestorbene: Deutschland, Jahre, Todesursachen
 Gestorbene: Deutschland, Jahre, Todesursachen
 GovData | May 22, 2018
 CSV XLSX

Krankheitskosten: Deutschland, Jahre, Krankheitsdiagnosen (ICD-10)
 Krankheitskosten: Deutschland, Jahre, Krankheitsdiagnosen (ICD-10)
 GovData | May 22, 2018
 CSV XLSX

Causes of death: Died, sex, age group (23), discarded cause...
 Causes of death: Died, sex, age group (23), discarded cause of death (European Short List, 80), Saxony, years
 GovData | May 20, 2018
 XLSX CSV

Disability (Physical and Sensory)
 The numbers of registered persons with physical or sensory disabilities classified by the Health Service Executive (hse) area and local health office, age group, sex, ...
 Data.gov.ie | March 9, 2018
 PDF

Stillbirths Maternal, Infant mortality and mortality

Countries

- Czech Republic (349813)
- Germany (202086)
- France (46408)
- United Kingdom (45906)
- Italy (38005)
- Spain (28977)
- Austria (23193)
- Netherlands (23020)
- European Catalogues (20513)
- Belgium (8007)

Show More Countries

Catalogues

- Geoportal Czech Office for Surve... (191972)
- GDI-DE (188961)

Asthma and Chronic Obstructive Pulmonary Disease (COPD) in Canada, 2018

Report from the Canadian Chronic Disease Surveillance System

Table of contents

- Executive summary
- Introduction
 - Canadian Chronic Disease Surveillance System (CCDSS)
 - Chapter 1: Asthma
 - Introduction
 - Asthma in the Canadian population
 - CCDSS Case Definitions
 - Prevalence
 - Incidence
 - All-cause Mortality
 - Comorbidity
 - Discussion
 - Chapter 2: Chronic Obstructive Pulmonary Disease
 - Introduction
 - COPD in the Canadian Population
 - CCDSS Case Definitions
 - Prevalence
 - Incidence
 - All-cause Mortality
 - Comorbidity
 - Discussion
 - Chapter 3: A Final Word
 - Moving Forward
 - Appendix A: Chronic Respiratory Disease Surveillance Development through the Canadian Chronic Disease Surveillance System
 - Canadian Chronic Disease Surveillance System
 - Chronic Respiratory Disease Feasibility Studies
 - Limitations of the CCDSS
 - Appendix B: Case Definitions for Comorbid Conditions
 - Comorbid Diseases or Conditions

Download the alternative format (PDF Version - 1.2 MB)

Organization: Public Health Agency of Canada

Published: April 2018

Related Topics

- Asthma
- Chronic Obstructive Pulmonary Disease (COPD)

©2017 Google LLC, used with permission. Google and the Google logo are registered trademarks of Google LLC

<https://www.europeandataportal.eu/data/en/datasets?tags=asthma>

<https://www.canada.ca/en/public-health/services/publications/diseases-conditions/asthma-chronic-obstructive-pulmonary-disease-canada-2018.html>

Challenges

- Discoverability and Access: Metadata
- Format and Structure
- Data Management
- Data quality and Integrity
- Harmonizing and linking open data from various sources
- Technology

Open Data Tools and Services

- [Open Data Tools](#): Several tools to Explore, Publish and Share Public Datasets
- [Developers Toolkit: Code for Canada](#): Access tools and resources to help develop innovative and creative applications using open data.
- [Canada's Open Data Exchange](#): National initiative which connects those who need data with those who have data, develops tools to share data, provides training and networking opportunities, and influences policy and standards.
- [Open Data Impact Map](#): A public database of organizations that use open government data from around the world.
- [The World Bank Open Data](#): Provides all users with open access to World Bank data. Includes access to Time Series, Microdata and Geospatial data.

Support the Open Data Community

- Make your data open (within reason)
 - Online availability and Open license
 - Machine readable, structured format (e.g. Excel instead of image scan)
 - Open non-proprietary format (e.g. CSV)
 - Append to a URI – Uniform Resource Identifier – to make it findable
 - Provide context – Link data to other data
 - Tim Berners-Lee, 5 Star Scale
- Use and Advocate for Open Data
- Learn and Educate

Questions & Discussion

Contact Information:

Vivek Jadon

Data Specialist

Library Data Service

Mills Library, McMaster University

vivek@mcmaster.ca

905 525 9140 Ext 23848

Laura Banfield

Librarian

Health Sciences Library

McMaster University

banfie@mcmaster.ca

905 525 9140 Ext 20239