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 WWW 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.
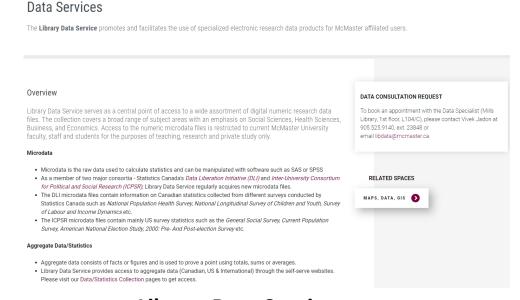
Opportunities

- 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

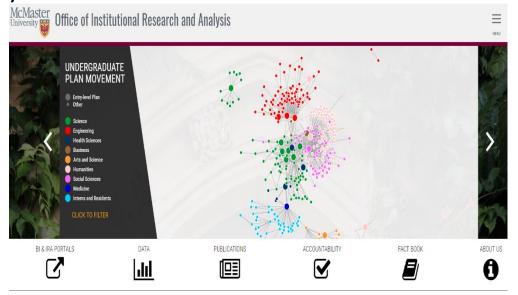- Education

- 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)

Library Data Service

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/
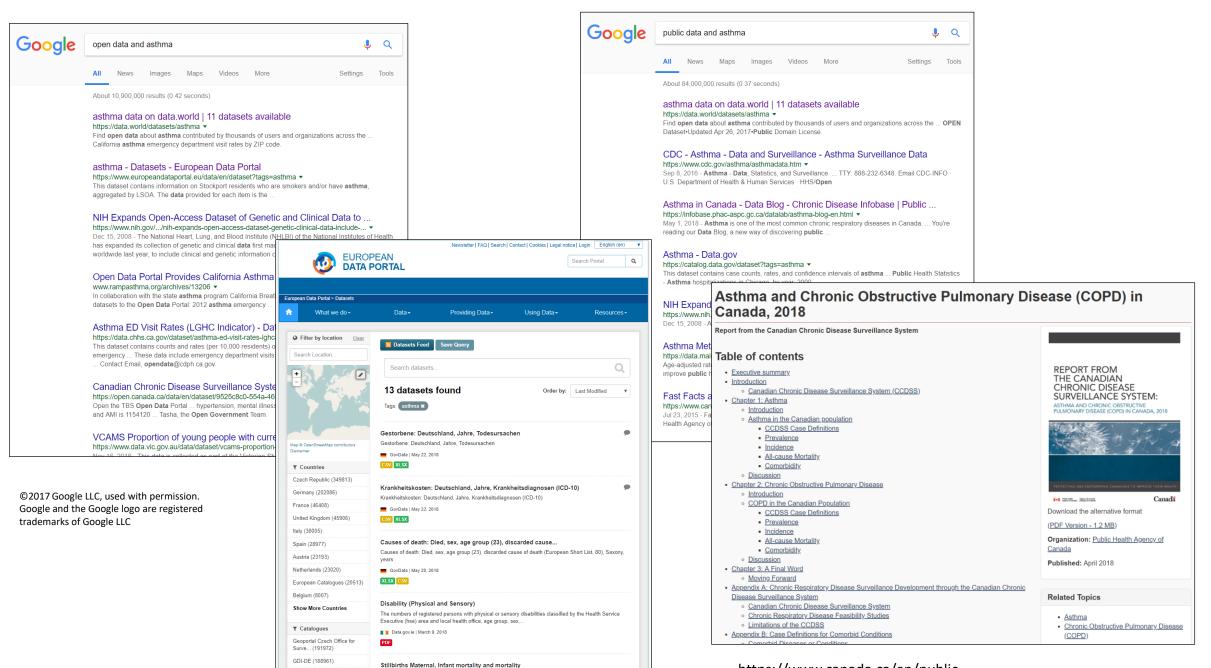- World Bank Open Data portal: https://data.worldbank.org/

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

https://www.europeandataportal.eu/data/en/dataset?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 uses with open access to World Bank data. Includes access to Time Series, Mircodata 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

http://opendatahandbook.org/

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