

# BCLibraries

### http://opendatahandbook.org/ en/what-is-open-data/

Liam Whalen - May 14, 2015

#### What is Open?

This handbook is about open data - but what exactly is *open* data? For our purposes, open data is as defined by the Open Definition:

Open data is data that can be freely used, reused and redistributed by anyone - subject only, at most, to the requirement to attribute and sharealike.

The full Open Definition gives precise details as to what this means. To summarize the most important:

- Availability and Access: the data must be available as a whole and at no more than a reasonable reproduction cost,
  preferably by downloading over the internet. The data must also be available in a convenient and modifiable form.
- Reuse and Redistribution: the data must be provided under terms that permit reuse and redistribution including the intermixing with other datasets.
- Universal Participation: everyone must be able to use, reuse and redistribute there should be no discrimination
  against fields of endeavour or against persons or groups. For example, 'non-commercial' restrictions that would
  prevent 'commercial' use, or restrictions of use for certain purposes (e.g. only in education), are not allowed.

If you're wondering why it is so important to be clear about what open means and why this definition is used, there's a simple answer: **interoperability**.

Interoperability denotes the ability of diverse systems and organizations to work together (inter-operate). In this case, it is the ability to interoperate - or intermix - different datasets.

Interoperability is important because it allows for different components to work together. This ability to componentize and to 'plug together' components is essential to building large, complex systems. Without interoperability this becomes near impossible — as evidenced in the most famous myth of the Tower of Babel where the (in)ability to communicate (to interoperate) resulted in the complete breakdown of the tower-building effort.

We face a similar situation with regard to data. The core of a "commons" of data (or code) is that one piece of "open" material contained therein can be freely intermixed with other "open" material. This interoperability is absolutely key to realizing the main practical benefits of "openness": the dramatically enhanced ability to combine different datasets together and thereby to develop more and better products and services (these benefits are discussed in more detail in the section on 'why' open data).

Providing a clear definition of openness ensures that when you get two open datasets from two different sources, you will be able to combine them together, and it ensures that we avoid our own 'tower of babel': lots of datasets but little or no ability to combine them together into the larger systems where the real value lies.

#### What Data are You Talking About?

What is Open? ■ What Data are You Talking About? How to Open up Data So I've Opened Up Some Data, Now Glossary Appendices Contribute Corrections and suggestions for improvements can be submitted on our issue tracker. If you are interested in contributing material please get in touch. Donate If you have found this useful and would like to support our work please consider making a small donation. - 2.00 US Dollar \$ Donate via PayPal Supported By

SHUTTLEWORTH



## http://opendefinition.org/licenses/

Liam Whalen - May 14, 2015

If you would like to propose a license to add to this page please follow our license approval process.

This section of the site lists licenses that are conformant with the principles laid out in the Open Definition.

#### **Conformant Licenses**

The following licenses are conformant with the principles set forth in the Open Definition.

- Domain = Domain of application, i.e. what type of material this license should/can be applied
  to. Note if you are looking for an open license for software, please see Open Source
  Definition conformant licenses.
- BY = requires attribution
- SA = require share-alike

#### Recommended conformant licenses

These licenses conform to the Open Definition and are:

- Reusable: Not specific to an organization or jurisdiction.
- Compatible: Must be compatible with at least one of GPL-3.0+, CC-BY-SA-4.0, and ODbL-1.0.
   Permissive/attribution-only licenses must be compatible with all 3 of the aforementioned licenses, and at least one of Apache-2.0, CC-BY-4.0, and ODC-BY-1.0.
- Current: Widely used and generally considered best practice by a broad spectrum of projects and actors within the domains of applicability of the license.

License	Domain	Ву	SA	Comments
Creative Commons CCZero (CCO)	Content, Data	N	N	Dedicate to the Public Domain (all rights waived)
Open Data Commons Public Domain Dedication and Licence (PDDL)	Data	N	N	Dedicate to the Public Domain (all rights waived)
Creative Commons Attribution 4.0 (CC-BY-4.0)	Content, Data	Y	N	
Open Data Commons Attribution License (ODC-BY)	Data	Y	N	Attribution for data(bases)
Creative Commons Attribution Share-Alike 4.0 (CC-BY-SA-4.0)	Content, Data	Y	Υ	
Open Data Commons Open Database	Data	_	_	Attribution-ShareAlike for

Français [Galego] пэээ [Croatian] Magyar [Indonesian | Islenska | Italiano | Јаралев | баја | 한국어 | македонски јазик | Norsk (bokmål) | Робизсухла | Ротидиёв Brasileiro | Ротидиёв | Русский [Shqip | Srpski | Svenska |



## https://opendata.dev4.catalogue.libraries.coop:4433/index.html

Liam Whalen - May 14, 2015

```
"titleO": "a Bowel / The Unitkely pligHings of Harold Fry ",
"earbeed": Joyce, Rachel.",
"prices": "Si3.95"
"subject heading": Finchion, Voyages and travels",
"subject heading": Finchion in Spouses",
"subject heading": Transchiptignion. "
"subject heading": "Recention. "
"subject heading": Tweenerflion. "
"subjec
```



### Generating Open Data

Liam Whalen - May 14, 2015

#### Sample SQL Query

SELECT bre.id, ac.circ\_lib, COUNT(acn.id) AS count FROM biblio.record\_entry AS bre INNER JOIN asset.call\_number AS acn ON bre.id = acn.record INNER JOIN asset.copy AS acp ON acn.id = acp.call\_number INNER JOIN action.circulation AS ac ON acp.id = ac.target\_copy INNER JOIN metabib.record\_attr AS mra ON bre.id = mra.id WHERE ac.circ\_lib IN (1301) AND (xact\_finish IS NULL OR xact\_finish > NOW() - interval '1 year') AND bre.deleted = false AND mra.attrs @> "search\_format"=>"physicalbooks" GROUP BY ac.circ\_lib, acn.id, bre.id ORDER BY ac.circ\_lib, count DESC LIMIT 100;

#### GIT Repo For Files Needed to Create the Open Data

http://git.sitka.bclibraries.ca/gitweb/?p=sitka/opendata-eg2015.git;a=summary



## http://geocoder.ca/?freedata=1

#### Liam Whalen - May 14, 2015

This data is provided under the Open Database License (ODbL) You haven't made it until you get sued Donate to geocoder.ca's legal defense fund to keep our data free

Submit your email address:

Submit Query

if you wish to be notified of future additions and updates to these free datasets.

This Geocoder.ca Dataset is made available under the Open Database License: http://opendatacommons.org/licenses/odbl/1.0/. Any rights in individual contents of the database are licensed under the Database Contents License: http://opendatacommons.org/licenses/dbcl/1.0/. Here at Geocoder.ca we provide a number of free geocoded data downloads, to help the community build better location based services. Check this page often for new additions to our free data collection. If you have a request for a geocoded dataset that is not included in this list, send use a note.

We are serious about providing reliable geocoding business services. These free data downloads showcase the high quality of our geocoding engine and data.

Your Free Data Downloads

#### **Geocoded IP Address List**

File of crowdsourced list of geocoded ip addresses. Includes ip address, latitude, longitude (Last Updated: April 7th 2014):

Geolocated IP Address List (687963 records)

#### **Canadian Postal Codes**

File of all Canadian postal codes. Includes postal, city, province, latitude, longitude (Last Updated: May 23rd 2014):

Canadian Postal Codes (918585)

#### US Zip Codes

File of all US 5-digit ZIP codes. Includes zip, city, state, latitude, longitude, and county (Last Updated: March 30th 2010):

• US Zip Codes (41,755)

#### **City Centroids**

We have a file for both Canada and the US which includes city name, state/province, latitude, and longitude (Last Updated: January 1, 2010):

- Canadian Cities (12,062 cities)
- US Cities (192,773 cities)

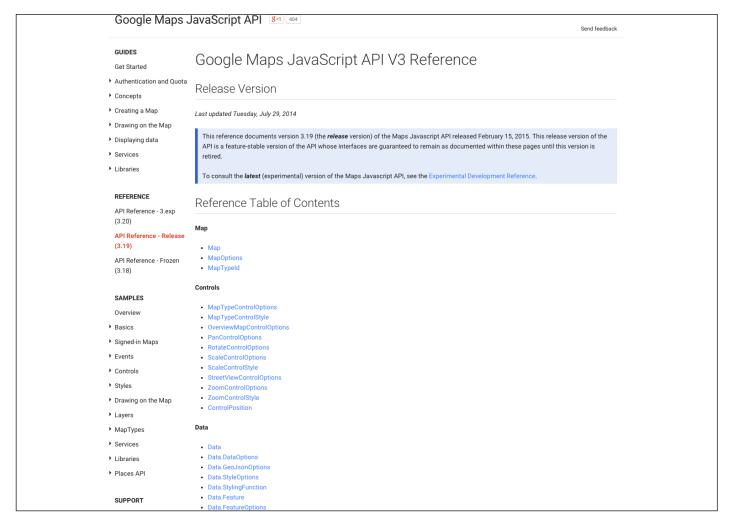
#### **Other Downloads**

- Crowdsourced Street address file 12232909 records. (for the past 6 months) As of November 7, 2014
- Crowdsourced Canadian Postal Code Polygon File As of May 5, 2013.



# https://developers.google.com/maps/documentation/javascript/reference

#### Liam Whalen - May 14, 2015





## Converting ArcGIS to GeoJSON OpenData

Liam Whalen - May 14, 2015

The data sets converted to GeoJSON were marked public at DataBC

DataBC at www.data.gov.bc.ca

Note the output file comes before the input file in the syntax

./ogr2ogr.py -f geoJSON geojson/prov\_lib\_data.json ./shpfiles/ BC\_Public\_Library\_Systems\_Locations\_and\_Branch\_Data.shp

The American Context

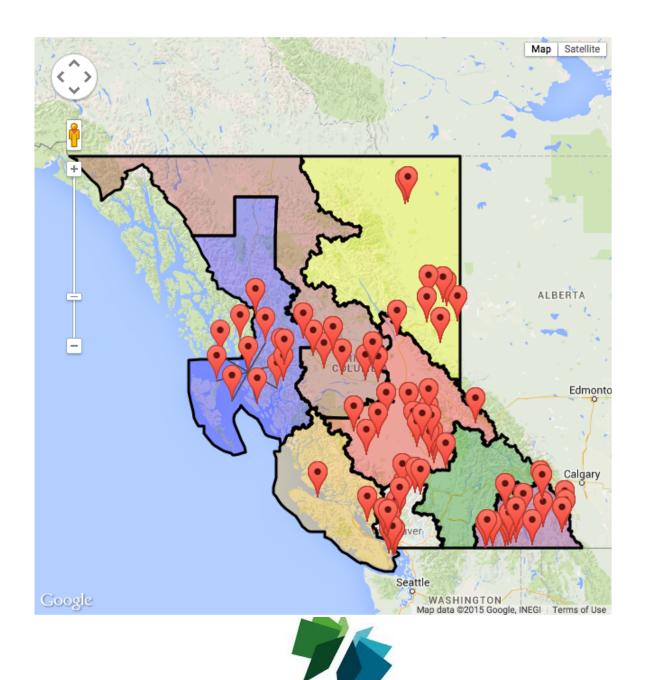
data.gov at <a href="https://www.data.gov/">https://www.data.gov/</a>

This site is packed with tons of data, and from a casual browsing the vast majority of it is available in Open formats.



## Library Book Data

Liam Whalen - May 14, 2015



BCLibraries COOPERATIVE

### Library Book Data

Liam Whalen - May 14, 2015

Combination of generated GeoJSON and 2011 Census Economic Regions data from DataBC.

Uses the Spiderfy javascript library to allow the Markers to radiate outwards allowing the display of the 8 most highly lent books per library.

https://github.com/jawj/OverlappingMarkerSpiderfier

GeoJSON automatically loaded by Google Maps' API.



## Average Price / Per Average Lend Heat Map

