



Statistics
Canada

Statistique
Canada

Canada



Statistics Canada
www.statcan.gc.ca



Statistics Canada's Dissemination Model: no semantic statistics?

Presented by:

Bill Joyce

October 2015





Outline

- Statistics Canada and Open Data
 - what we are doing
- Statistics Canada and Open Data
 - what we are not doing
 - why not?
 - are we being wise and prudent or are we missing the boat?

Statistics Canada is adopting open data principles

- The concept of open data is a natural fit with Statistics Canada's mission
- Over the years there has been an evolution toward free data on the Statistics Canada website
- Most data is downloadable (CSV and SDMX-ML)
- Since February 2012
 - All standard aggregate data on the Web is **free of cost**
 - Adoption of an **Open Data Licence**
 - **Elimination** of licensing and **royalty fees**, even for custom data



The Government of Canada Open Data (Open Government) Portal

Open Data Portal

Search Datasets Suggest a Dataset

244,602 datasets found Order by Last Modified

National, Provincial and Territorial Air Pollutant Emission Inventory (APEI),...

Canada's Air Pollutant Emission Inventory (APEI) is a comprehensive inventory of air pollutant emissions at the national, provincial and territorial levels, prepared and...

Organization: Environment Canada

Resource Formats: **XLS** HTML

Air Pollutant Emission Inventory (APEI) Historical Trends, 1990-2013

Search Filters

Organization

- Agriculture and Agri-Food Canada (1821)
- Department of Finance Canada (130)
- Environment Canada (247)
- Natural Resources Canada (235913)
- Statistics Canada (5995)

▼ Show more

Data Type

- Raw Data (8140)
- Geo Data (236462)

75%

- StatCan is also the service provider responsible for the government wide portal
 - Common CKAN technology for our inventory and for open.canada.ca



Coming soon

- Static data publications will evolve toward the provision of data in a more open database-driven model
- We will introduce more consistent file downloads and web data services



What we are not doing

Survey of household spending (SHS), household spending, by household tenure

Survey of household spending (SHS), household spending, by household tenure

Licence:
Open Government Licence - Canada

Dataset Resources

Resource Name	Format	Language	Link
Dataset	CSV	English	Download
Dataset	CSV	French	Download
Dataset	XML	Bilingual (English and French)	Download
Supporting Document	HTML	English	Download
Supporting Document	HTML	French	Download

Developer Tools
The information on this page (the dataset metadata) is also available in JSON format
[Link to JSON format](#)

Rate this dataset
0 Comment(s)

[About this Dataset](#)
[RSS Feed](#)

Publisher:
Statistics Canada

Tags:
CANSIM

Subject:
Economics and Industry

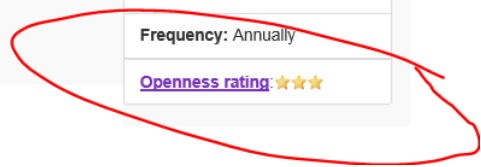
Date Published:
2015-01-22

Date Modified:

Program Page URL: [Program Page](#)

Frequency: Annually

Openness rating: ★★☆☆





Semantic Statistics – why not?

- Other priorities
 - A new dissemination model
 - simplified and streamlined product line
 - single consolidated output database - including WDS
 - single comprehensive inventory and improved navigation



Semantic Statistics – why not?

- No indication from any data users that this is a requirement
 - and we have been asking our data users to tell us what they need in terms of output
 - two rounds of focus groups
 - a beta test with on line feedback

- Is the work towards semantic statistics somewhat theoretical and academic?



Semantic Statistics – why not?

- Will the concept really take off?
- Even if it does, will it take off in its current form?
- Or, will a linked data format evolve to a point where it is substantially different from its current form?
- Does it makes sense, from a selfish point of view, to wait until the concept matures and settles, until there is a well vetted recipe book?

Semantic Statistics – why not?

- Even if the concept is mature enough now, and the user need is real, would it make any sense to leave the heavy lifting to a third party?
- We can't be all things to all people, we need to be the supplier of data, and, as in the case with mobile applications and sophisticated data visualizations, could there be a role for third parties to take our data and add value by formatting it for the semantic web?

Semantic Statistics – are we 'missing the boat' ?

- OR... will we realize, at a certain point in time, that a new generation of data users will be consuming statistics only via Siri and Watson?
- If our data, as we make it available, is not formatted adequately, do we risk losing our position as the authoritative source of official statistics for Canada?
- If this is the case, and if that day is near, where the heck do we catch this boat?



Questions or comments?

Bill Joyce
Director
Dissemination Division
Statistics Canada
(613) 951-2683
bill.joyce@canada.ca