

DIALLED

Distributed Index of *All* *some* Library Location and Event Data

Dan Scott, January 29, 2016





Allana Mayer

@allanaaaaaaa



Following

... wait, did I just read the entire OLA schedule and not see a single thing about linked open data

LIKES

3



6:05 PM - 25 Jan 2016



Dan Scott @denials · Jan 25

@allanaaaaaaa The payoff from this is a Turtle file, so hopefully that counts?
olasuperconference.ca/event/dialled-...



1



[View summary](#)

The 1-minute version

- There is linked data!
- Current data dump in Turtle format is at <https://dialled.ca>
- There is not a huge amount of data
- There is still a lot of work to do
- Budget = my time + \$0
- Please help (both lightweight and heavyweight participation is encouraged!)

Canada's Libraries and Archives

Expert Panel on *The Status and Future of Canada's Libraries and Archives*

Published: June 2014

Searching: Search below by typing keywords, and/or limiting using the various filters that are available. Filtering options include establishment year, institution type, type of library, and province/territory. To restart your search, click on "Reset All Filters", which will appear above the map after keywords and/or limiting filters have been selected.

Keywords:

Establishment year range

789

Institutions

Archives

Combined

Types

Academic libraries

Archives

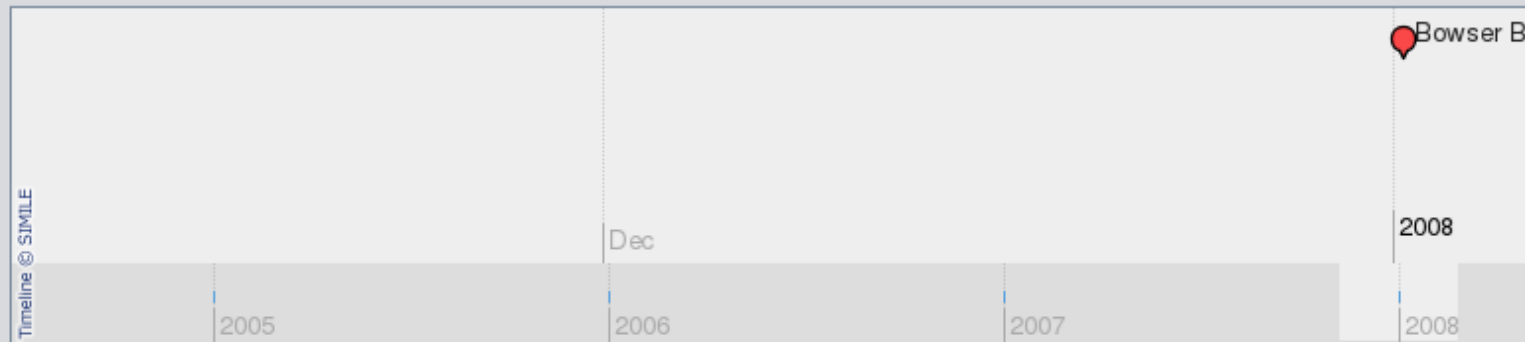
Provinces/Territories

Alberta

British Columbia

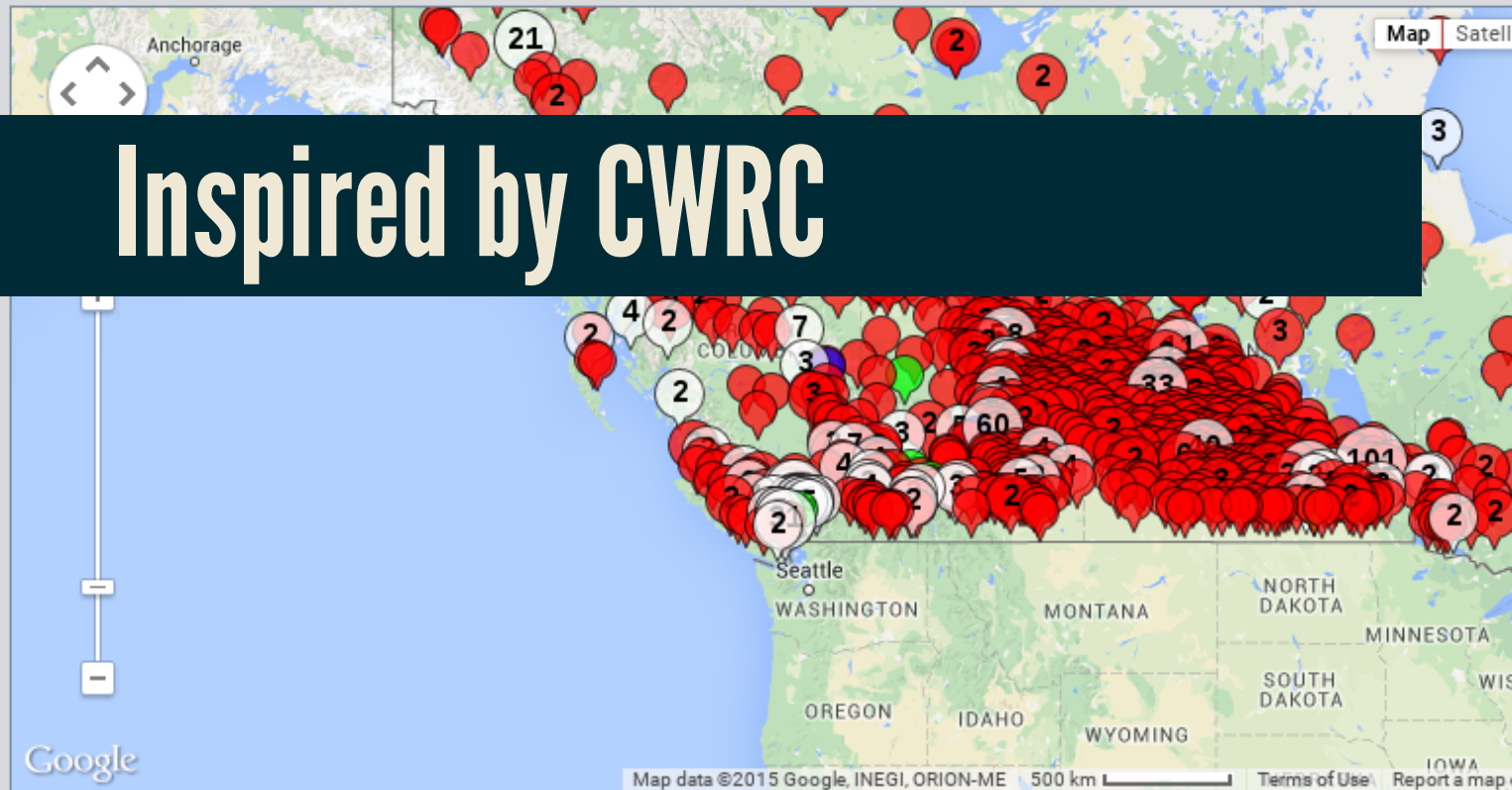
Background: This mapping and timeline application displays the location of, and the year of establishment of many of the libraries and archives of Canada. Both libraries and archives are represented, and for institutions that function as both a library and an archive, they are indicated as "Combined" institutions. "Mixed" indicates a particular community has more than one type of institution. All 4,798 represented institutions appear on the map, and roughly two-thirds of the institutions (3,077 of 4,798, or 64%) appear on the map.

1721 results out of 4798 cannot be plotted

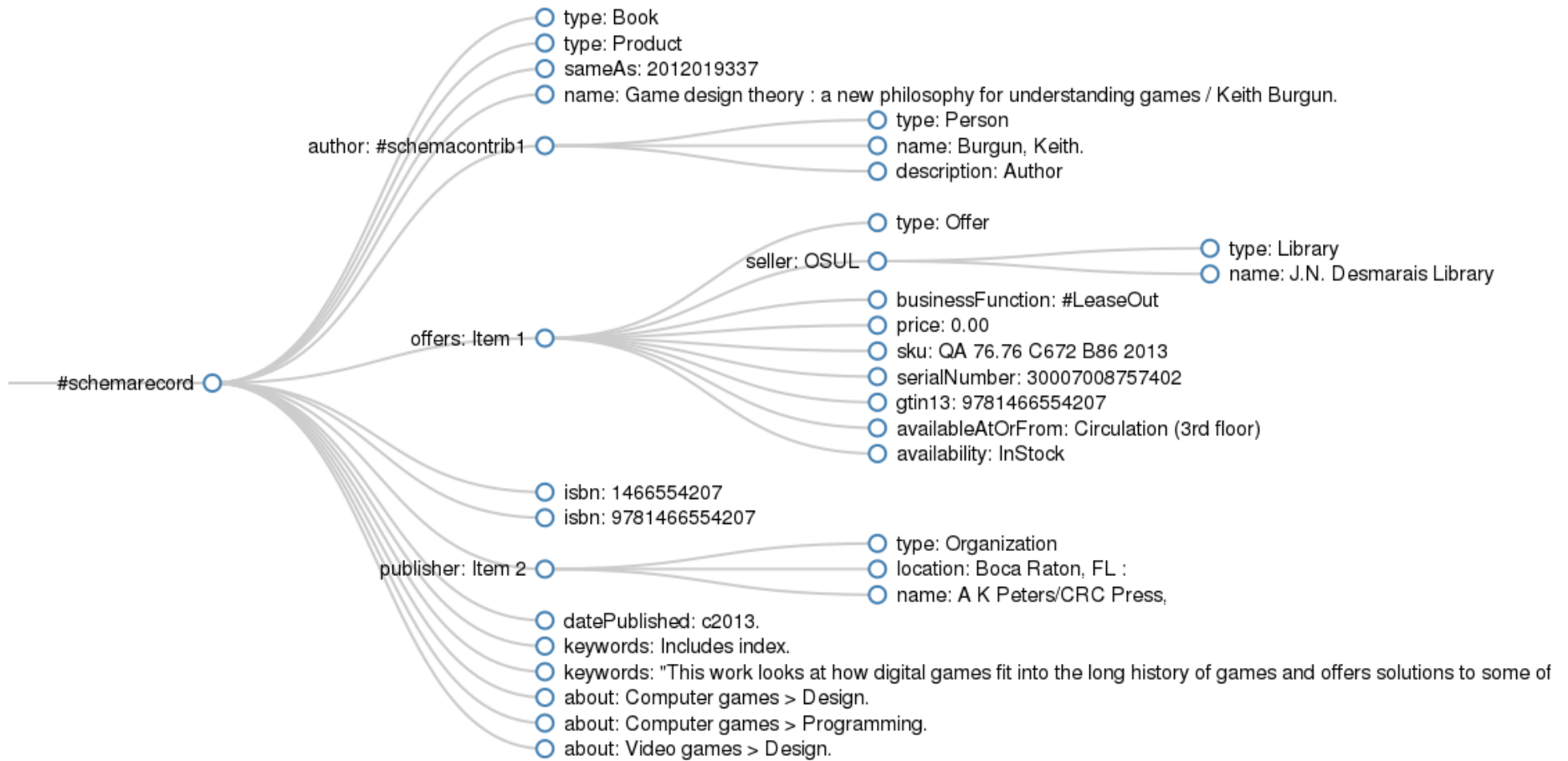


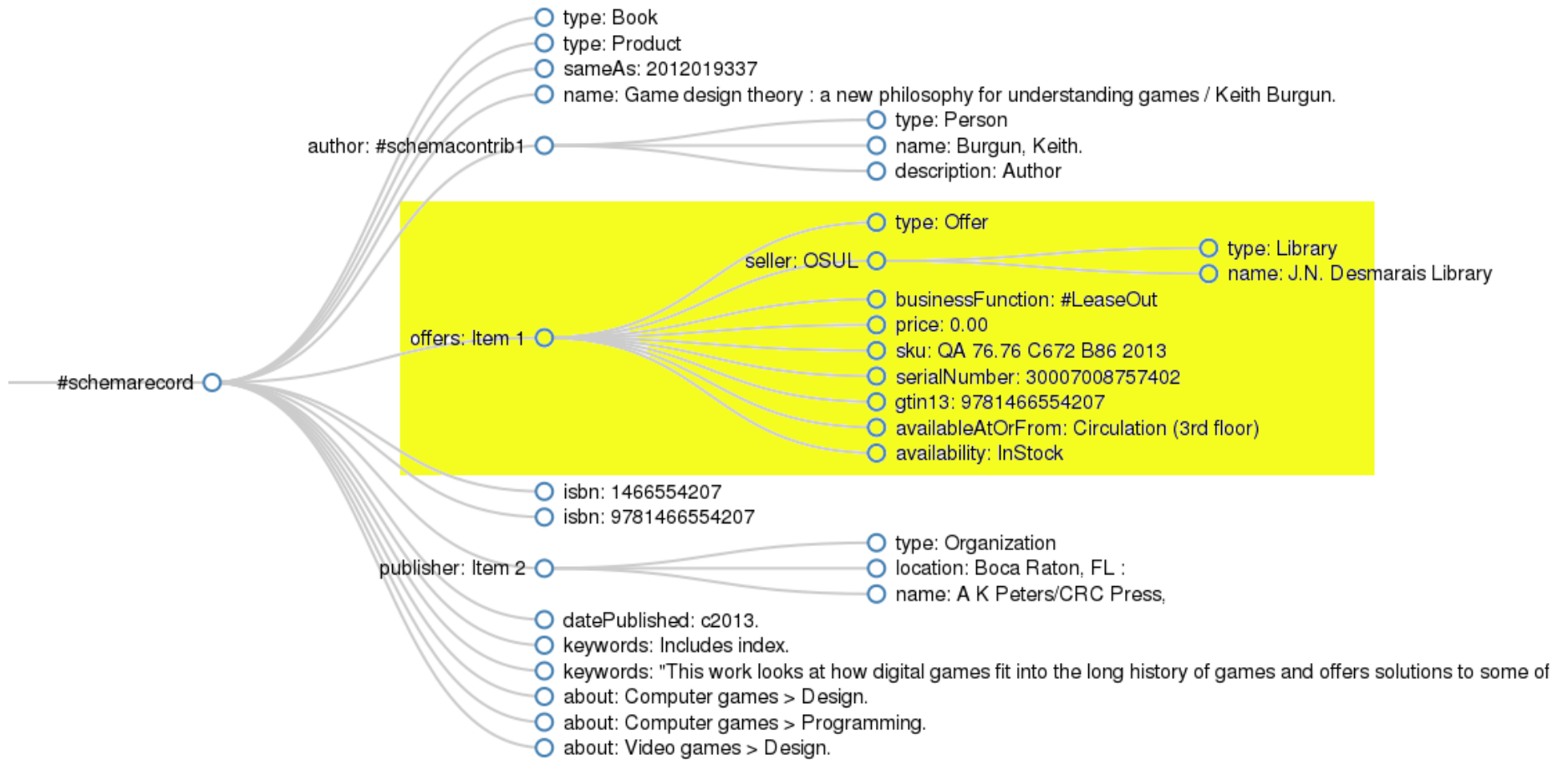
MAP VIEW • LIST VIEW • GRID VIEW

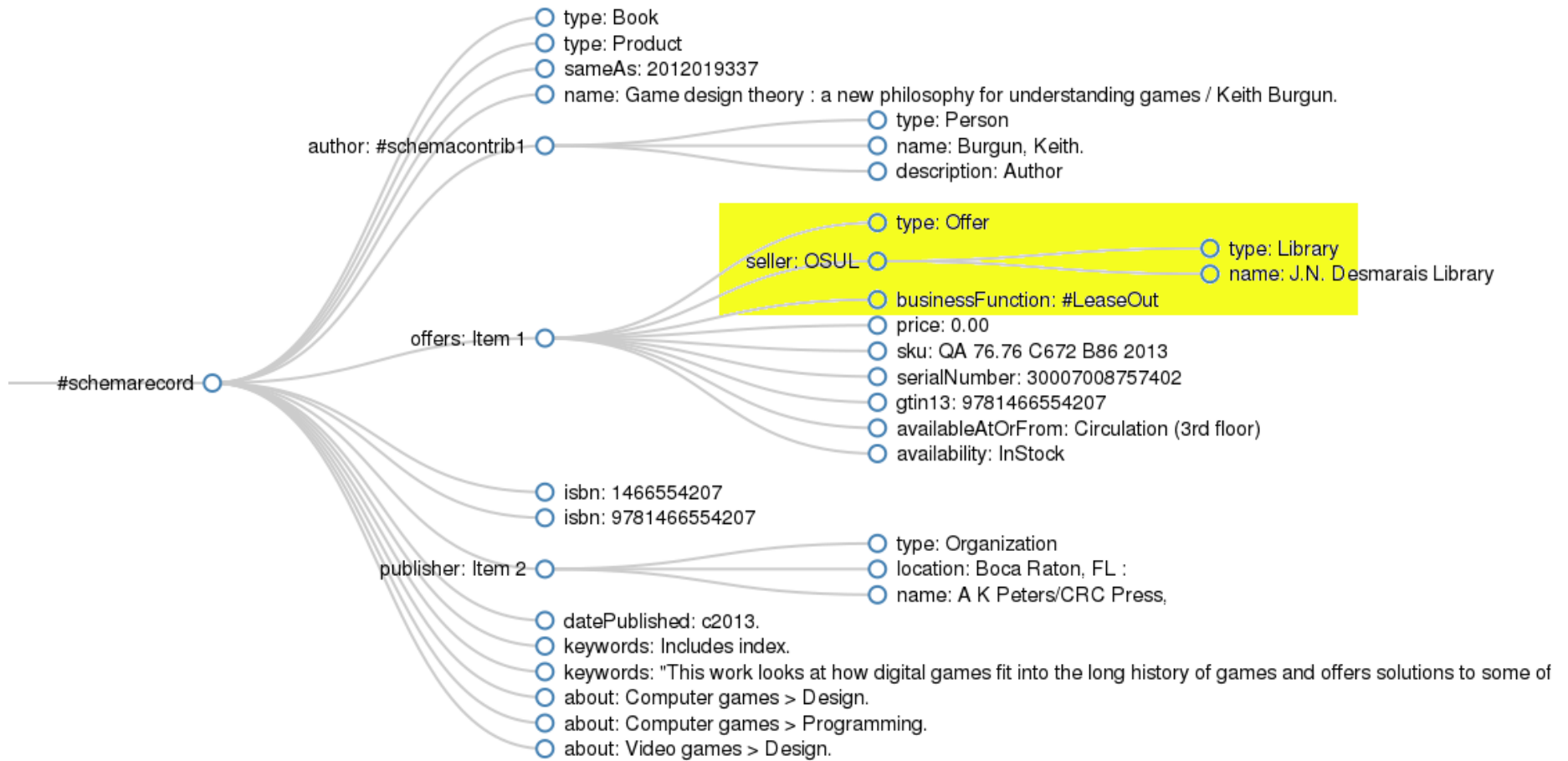
4798 items



Archives Combined Libraries mixed







Library linked data principles

- Every bibliographically described resource, its instances
- Every instance, its offering library
- Every library, its URL!
- ... with hours, location, contact data so people can get to your stuff

I started at the bottom

- Our MARC records now link to OCLC Work URIs and LoC IDs
 - Evergreen expresses these using schema.org in RDFa
- OCLS is similarly enriching their MARC records
- Zepheira offers a MARC->BIBFRAME->HTML + schema.org service

WorldCat Registry

authoritative single source for institutional metadata

- Problematic non-commercial license
- Mostly missing or out-of-date data

Library and Archives Canada



Library and Archives
Canada

Bibliothèque et Archives
Canada

Canada



Library and Archives Canada

www.collectionscanada.gc.ca

[Français](#)

[Home](#)

[Contact Us](#)

[Help](#)

[Search](#)

[canada.gc.ca](#)

[Home](#) > [Libraries](#) > [Union Catalogue](#)

Union Catalogue

Results

[Search](#) | [Search Help](#)

Search Term(s): "Toronto" and "Ontario" and "Public"

Results: (1-4 of 4)

Library Symbol	Library Type	Library Name
1. QTYP	Public	City of York Public Library, Toronto
2. QTEY	Public	East York Public Library, Toronto
3. QIP	Public	Toronto Public Library, Bibliographic Services, Toronto
4. QTMCL	Public	Toronto Public Library, Toronto

Results: (1-4 of 4)

[Social Tagging](#) ([About Social Tagging](#))

[RSS](#) [RSS Feeds from LAC](#)

[Add to Delicious](#) [Add to Digg](#) [Add to Diigo](#) [Add to facebook](#) [Add to Technorati](#)

Date Created: 2009-01-15
Date Modified: 2009-01-15

[Top of Page](#)

[Terms and Conditions](#)

Germans rock linked open data

- beta.lobid.org offers a web API for a comprehensive dataset of library institutions in Germany
- Major hat tip to Adrian Pohl!

```
GET http://beta.lobid.org/organisations/DE-6
{
  "isil" : "DE-6",
  "wikipedia" : "http://de.wikipedia.org/wiki/Universitäts-_und_Landesbi",
  "@context" : "http://beta.lobid.org/organisations/context.jsonld",
  "linkedTo" : "http://lobid.org/organisation/DE-605",
  "@type" : "http://schema.org/Library",
  "rs" : "055150000000",
  "name" : "Universitäts- und Landesbibliothek Münster",
  "url" : "http://www.uni-muenster.de/ULB/",
  "provides" : "http://www.uni-muenster.de/ULB/katalog/",
  "@id" : "http://beta.lobid.org/organisations/DE-6#!",
  "telephone" : "0049 251 83-24040",
  "email" : "mailto:info.ulb@uni-muenster.de",
  "address" : {
    "postOfficeBoxNumber" : "80 29",
    "addressLocality" : "Münster",
    "@type" : "http://schema.org/PostalAddress",
```

So, about that CWRC work...

Canada's Libraries and Archives

RSC Expert Panel on *The Status and Future of Canada's Libraries and Archives*

Published: June 2014

Searching: Search below by typing keywords, and/or by limiting using the various filters that are available. Limiting options include establishment year, institution type, type of library, and province/territory. To restart your search, click on "Reset All Filters", which will appear above the map after keywords and/or limiting filters have been selected.

Keywords:

Establishment year range

 -

Institutions

518 Archives

42 Combined

Types

428 Academic libraries

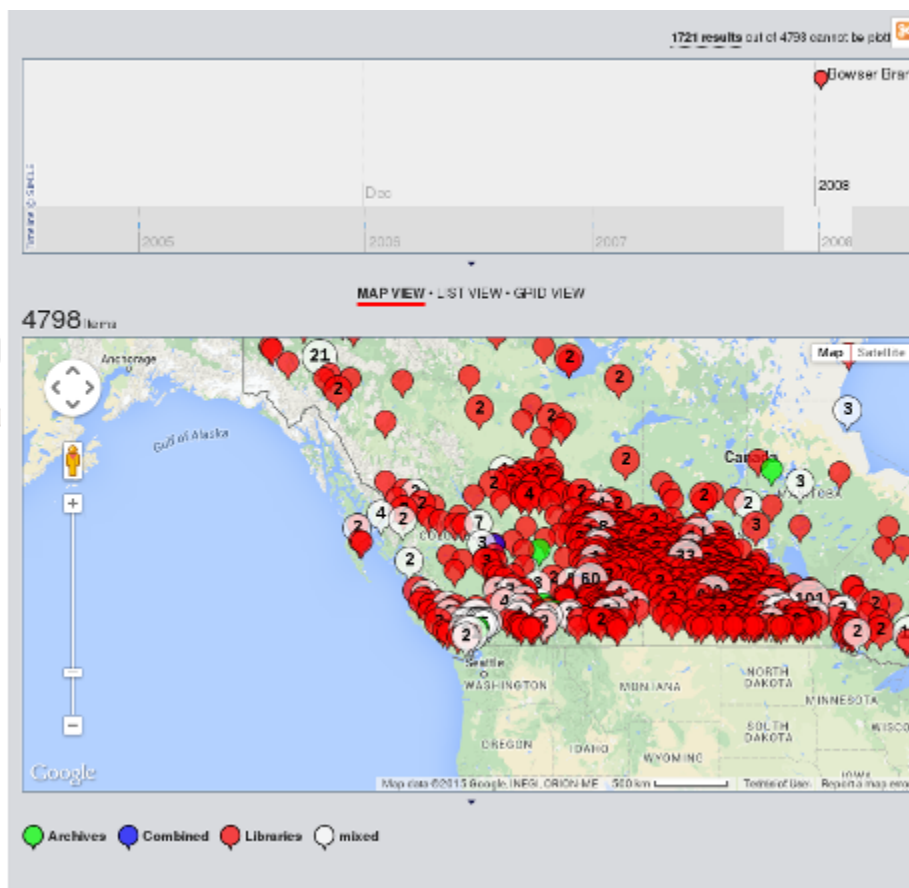
518 Archives

Provinces/Territories

442 Alberta

424 British Columbia

Background: This mapping and timeline application displays the location of, and the year of establishment for, many of the libraries and archives of Canada. Both libraries and archives are represented, and for institutions that function as both a library and an archive, they are indicated as "Combined" institutions ("mixed" indicates a particular community has more than one type of institution). All 4,798 represented institutions appear on the map, and roughly two-thirds of the institutions (3,077 of 4,798, or 64%) appear on



CWRC's dataset

```
{  
  "community": "Sudbury",  
  "group": "Libraries",  
  "institutionType": "Library",  
  "label": "Laurentian University",  
  "latLng": "46.5011111,-80.9775",  
  "latitude": "46.5011111",  
  "libraryType": "Academic",  
  "longitude": "-80.9775",  
  "population": null,  
  "provCode": "ON",  
  "province": "Ontario",  
  "sort": "707",  
  "startDate": "1960",  
  "subGroup": "Academic libraries",  
  "url": "http://library.laurentian.ca/content/library"  
}
```

Step 1. Crawl the URLs

- 4,798 entries in the CWRC dataset
- 4,250 with URLs
- 3,223 unique URLs
- 561 URLs redirected
- 3,141 URLs retrieved

Step 2. Update the CWRC dataset

```
{  
  "...": "...",  
  "url": "http://laurentian.ca/library",  
  "urlPrevious": "http://library.laurentian.ca/content/library"  
}
```


Step 3. Extract structured (or linked) data

- Microdata (`itemscope`, `itemtype`, `itemprop`): 191 pages
- RDFa (`vocab`, `typeof`, `property`): 197 pages
- JSON-LD (`<script type="application/ld+json">`): 112 pages

Encouraging!

Step 4. Evaluate the extracted data

Some of it is weaksauce...

```
<a href="/node/1486">  
    
</a>
```

Step 4. Evaluate the extracted data

Some of it rocks!

```
<script type="application/ld+json"> {  
  "@context" : "http://schema.org",  
  "@type" : "Organization",  
  "name" : "Viking Municipal Library ",  
  "url" : "http://www.vikinglibrary.ab.ca",  
  "contactPoint" : {  
    "@type" : "ContactPoint",  
    "contactType" : "Customer Service",  
    "telephone" : "+(780) 336-4992",  
    "faxNumber" : "+(780) 336-4992",  
    "areaServed" : "CA",  
    "availableLanguage" : "English",  
    "email" : "mailto:librarian@vikinglibrary.ab.ca"  
  },  
  "hasPOS" : {  
    "@type" : "Place",  
    "name" : "Viking Municipal Library",  
    "hasMap" : {  

```

Step 5. Publish the linked open data

One big Turtle file

```
<http://biblio.uottawa.ca/en/morisset-library> md:item ( [ a schema:Library ;
  schema:branchOf [ a schema:Library ;
    schema:address [ a schema:PostalAddress ;
      schema:addressCountry "Canada"@en ;
      schema:addressLocality "Ottawa"@en ;
      schema:addressRegion "ON"@en ;
      schema:postalCode "K1N 6N5"@en ;
      schema:streetAddress "65 University Private"@en ;
    schema:alternateName "Bibliothèque Morisset"@en ;
    schema:geo [ a schema:GeoCoordinates ;
      schema:latitude "45.423241"@en ;
      schema:longitude "-75.684126"@en ] ;
    schema:name "Morisset Library"@en ;
    schema:url <https://biblio.uottawa.ca/en/morisset-library> ;
    schema:name "uOttawa Library"@en ] ) ;
rdfa:usesVocabulary schema: .
```

The picture in Canada

- schema.org/Library *Usage*: 3 libraries--one of which is an archives
- schema.org/LocalBusiness *Usage*: 47 statements
- schema.org/Organization : *Usage*: 15 statements
- schema.org/Event : *Usage*: 44 statements
- schema.org/PostalAddress : *Usage*: 103 statements

This should be in our wheelhouse!

Why should you care--really?

- **Visibility and accuracy** in common search engines
- Ease of building lightweight applications
- General linked open data use cases

What can you do?

- Add linked data to your library home page
- Help manage the list of libraries
- Help build dialled.ca

Adding linked data

Scott, D. (2015). White Hat Search Engine Optimization (SEO): Structured Web Data for Libraries. *Partnership: The Canadian Journal of Library and Information Practice and Research*, 10(1). Retrieved from <https://goo.gl/uMclgE>


Web page before linked data

Niceville Library

[Log in/Sign up](#)
[Location/Hours](#)

[Catalog](#) Search books and more


- Books & More
- Events
- For Kids
- For Teens
- Become a Member
- Research
- Other Services
- About Us



Location/Contact **Hours**

(555) 555-5555
hello@nicevillelibrary.org

Monday-Friday: 10:00 AM-9:00 PM
Saturday & Sunday: Noon-9:00 PM



[Niceville Library](#)
[143 Lucky Hi-way](#)
[Niceville, TN](#)
[USA 12345](#)




Web page after linked data

Niceville Library

[Log in/Sign up](#)
[Location/Hours](#)

[Catalog](#) Search books and more


- Books & More
- Events
- For Kids
- For Teens
- Become a Member
- Research
- Other Services
- About Us




Location/Contact **Hours**

(555) 555-5555
hello@nicevillelibrary.org

Monday-Friday: 10:00 AM-9:00 PM
Saturday & Sunday: Noon-9:00 PM



[Niceville Library](#)
[143 Lucky Hi-way](#)
[Niceville, TN](#)
[USA 12345](#)



Education by demonstration!

- An augmented Library One-Pager
- Live at dialled.ca/one_pager
- git branch at github.com/dbs/One-Pager/tree/schema.org

Standard RDFa tactics

```
@@ -14,14 +14,14 @@
    <link rel="stylesheet" href="deps/ie.css">
    <script src="//html5shiv.googlecode.com/svn/trunk/html5.js"></script>
    <![endif]->
- </head><body><div class="page">
+ </head><body vocab="http://schema.org/" typeof="Library"><div class="page">

  <header>
    <nav>
      <a href="#">Log in/Sign up</a>
      <a href="#footer">Location/Hours</a>
    </nav>
- <h1><a href="#">Niceville Library</a></h1>
+ <h1 property="name"><a href="#">Niceville Library</a></h1>
  <form id="search"><fieldset><legend>Search Catalog</legend>
    <label for="searchText"><a href="#">Catalog</a></label><input type="submit" value="Search">
  </form>
</div></body></html>
```

Reality check

- How many of you use a static HTML library web site?
- How many use a CMS like Drupal?
- How many of you want a deep knowledge of RDFa, microdata, or JSON-LD and HTML to be the minimum bar for someone to change the opening hours on your website?

Drupal One-Pager

- A theme with user-friendly widgets that hide any implementation details.
- So easy, a University Librarian could use it!™

Location and contact info

ere **Edit Library Address Location/Contact** 

[Home](#) » [Location/Contact](#)

Title *

Location/Contact

Email

nicevillelibrary@example.org

Telephone

(555) 555-5555

ADDRESS

Country

Canada 

Address 1 *

935 Ramsey Lake Road

Address 2

City *

Sudbury

Province *

Ontario 

Postal code *

P3E 2C6

This really isn't hard!

- **NO WEBSITE REDESIGN NECESSARY**
- Nobody needs to know except you and the world of linked open data

How to help dialled.ca directly

- Ensure your own library is listed, with an accurate URL: see <https://goo.gl/ZpsB8M>
- Create and maintain machine-readable lists
- Write HTML scrapers for existing human-oriented lists

The future of dialled.ca

- Generate periodic data dumps and statistics of the crawled data
- Normalize the data across different vocabularies
- Add more data from more sources (ILL symbols, etc)
- Integrate checks and updates via external lists
- Offer a dynamic web interface for querying the existing data
- Enhance the dialled.ca user experience
- Contributions welcome at <https://github.com/dbs/dialled-crawler>