Building a web based health data search tool using DDI

Amber Leahey
Building a web based health data search tool using DDI

Amber Leahey
Project Partners

Population Health Improvement Research Network

Ontario Council of University Libraries
Data Knowledge Sharing

Disparate sources

- Canadian Institute for Health Information
- Ontario Ministry of Health and Long-Term Care
- Ontario Ministry of the Environment
- ICES
- OMA

Open Access Curriculum

- Education
- Health Professionals
- Lots of Data!
- Laboratories

Health care
Hospitals
Disparate sources
Managing access information / licensing around health data is exhausting!

Open

“A piece of data or content is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike.” — See more at: [http://opendefinition.org/mbahn/pl/3ykT1I4upd/](http://opendefinition.org/mbahn/pl/3ykT1I4upd/)

Example:

“Cop, modify, publish, translate, adapt, distribute or otherwise use the data in any medium, mode or format for any lawful purpose with acknowledgment.” — See more at: [https://ontario.ontario.ca/government/open-government-license-ontario](https://ontario.ontario.ca/government/open-government-license-ontario)

Open - ish (free vs. open)

- Ongoing/Available
- "Contact us for more information"
- CD-ROM, DVD, XML, proprietary formats
- Format and usability issues
- Time delays (email, mail delivery etc.)
- Cost recovery for delivery / transfer of data

Mediated Access

- DLJ
  - The Data Liberation Initiative (DLI) is a partnership between the Government of Canada and several universities and research institutions. It provides access to confidential data for research and analytics.
  - The data is hosted on a secure platform and can be accessed through a secure web interface.
  - The data is made available to researchers and organizations who are approved by the Government of Canada.

- Open Data Program
  - The Open Data Program provides access to government data in a machine-readable format.

- Local Information Details (LID)
  - The Local Information Details (LID) program provides access to local government data.

- CHIN Graduate Student Data Access Program (GSDAP)
  - Access to data for graduate students for research purposes.

Cost Recovery Programs

- CIHI Data
  - The Canadian Institute for Health Information (CIHI) provides data and analytics for health care providers, hospitals, and other organizations.

- Cost Recovery: Statistics Canada
  - The Statistics Canada Cost Recovery program provides access to data for government departments and non-profit organizations.

Restricted Access Services

- CIHI Portal
  - Access to data for health care providers, hospitals, and other organizations.

- Statistics Canada Real-Time Remote Access (RTRA)
  - Access to data for government departments and non-profit organizations.

- Statistics Canada Research Data Centre (RDC)
  - Access to data for research organizations.

- Government Department Information Systems
  - Access to data for government departments and non-profit organizations.
Open

“A piece of data or content is open if anyone is free to use, reuse, and redistribute it — subject only, at most, to the requirement to attribute and/or share-alike.” - See more at: http://opendefinition.org/#sthash.FL3ylH1P.dpuf

Example:

"Copy, modify, publish, translate, adapt, distribute or otherwise use the data in any medium, mode or format for any lawful purpose with acknowledgment." - see more at: https://contrib.ontario.ca/government/open-government-licence-ontario
Open-ish (free vs. open)

- Ongoing/Available
- "Contact us for more information"
- CD-ROM, DVD, .XML, proprietary formats

- Format and usability issues
- Time delays (email, mail delivery etc.)
- Cost recovery for delivery / transfer of data
Mediated Access

DLI
“The Data Liberation Initiative (DLI) is a partnership between post secondary institutions and Statistics Canada for improving access to Canadian data resources. Over the years the focus of the DLI Program has evolved from purchasing access to major Canadian datasets collected by Statistics Canada...” See more at: http://www.statcan.gc.ca/dli-idd/dli-idd-eng.htm

Land Information Ontario (LIO)

CIHI's Graduate Student Data Access Program (GSDAP)

Municipal consortia
• Community Data Program
Cost Recovery Programs

ICES Data & Analytic Services (DAS)

CIHI Clients

"CIHI responds to data requests on a cost-recovery basis, as follows:

$145/hour, for Canadian health care facilities, media (data requests only), governments, not-for-profit health agencies, universities (except for students through the Graduate Student Data Access Program), health professionals and researchers from the public sector (Price A clients)

$290/hour for private commercial operations (including but not limited to software vendors and consultants), foreign clients (Price B clients) and others not qualifying for Price A.”

See more at: http://www.cihi.ca/C/II-lxt-portal/internet/EN/TabbedContent/standards+and+data+submission/data+requests/cihi012210

Statistics Canada

• Custom tabulations
Restricted Access Services

CIHI Portal
- health care providers, hospitals, regional health authorities

Statistics Canada's Real-Time Remote Access (RTRA)
- Government departments, non-profits, academics

Statistics Canada's Research Data Centres (RDCs)

Government Department Information Systems
DDI Codebook Tags

<copyright>, <distrbtr>, <contact>, <biblCit>, <specPerm>,
<restrctn>, <citReq>, <deposReq>, <conditions>,
<disclaimer>

<restrctn> 2.4.2.3 Restrictions

Optional
Not Repeatable
Attributes: ID, xml:lang, source

Description: Any restrictions on access to or use of
the collection such as privacy certification or
distribution restrictions should be indicated here.
These can be restrictions applied by the author,
producer, or disseminator of the data collection. If
the data are restricted to only a certain class of
user, specify which type.
Ontario Population Health Index of Databases (OPHID)

Search: ambulance in Anywhere

Refine By Date: Earliest to 2011

Need help getting started?
Making Data Reusable

Common Language
Accessibility
Commitment to open access

Sustainable infrastructure

Sustainable service model

DDI
Focus on research tools

Preservation and don't rely on government
Infrastructure & Sustainability

- Persistent Identifiers
  - URLs
  - DOIs?

- Expertise required
  - review data
  - create new records

- Growth required
  - not complete, ongoing collection

- Marketing
  - like with any tool, researchers need to know about it
Evaluate Research Benefit

- Google Analytics
- Researcher feedback
- Government feedback

<table>
<thead>
<tr>
<th>Does not</th>
<th>Does</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assign unique identifiers to data</td>
<td>Central place for researchers</td>
</tr>
<tr>
<td>Complete inventory of data sources</td>
<td>Common language</td>
</tr>
<tr>
<td>Cross reference data</td>
<td>Expose data gaps; access gaps</td>
</tr>
<tr>
<td>Provide access to data</td>
<td>Free, online</td>
</tr>
</tbody>
</table>

Still just another tool...

Bring together communities of people to talk about data!
Government feedback

Does not

- Assign unique identifiers to data
- Complete inventory of data sources
- Cross reference data
- Provide access to data

Still just another tool...

Does

- Central place for researchers
- Common language
- Expose data gaps; access gaps
- Free, online

Bring together communities of people to talk about data!
Building a web based health data search tool using DDI

Amber Leahey