Ontario Data Documentation, Extraction Service and Infrastructure

Tuesday May 31\textsuperscript{st}, 2011
DLI Boot Camp
Vancouver, British Columbia
Jane Fry (Carleton University)
Susan Mowers (University of Ottawa)
Overview

- <odesi> is ...
- History
- The Process
- Features
- Future Directions
<odesi> is ...

- centralised
- standardised
- bilingual
- a web-based data exploration system
- a web-based data extraction system
- delivered through Scholars Portal
- available to the 21 OCUL institutions (Ontario Council of University Libraries)
What is Scholars Portal?

- A service of OCUL

- Some of the other services offered
  - Interlibrary loan - RACER
  - Reference management - RefWorks
  - Digital content for Journals
  - Digital content for Books
  - Link server – SFX
  - Webinars on SP services

- www.scholarsportal.info/
History of <odesi>

- In 2005, DINO (Data in Ontario) started thinking about moving beyond ‘home-grown’ data solutions.
  - Jeff Moon (Queen’s)
  - Bo Wandschneider (Guelph)
  - Wendy Watkins (Carleton)

- Data Centres were interested in building a province-wide data solution.

- DDI was being recognised as a coding standard.
  - But we needed software to interpret DDI.
**History (cont’d)**

- **Nesstar** ([www.nesstar.com](http://www.nesstar.com))
  - Networked Social Science Tools and Resources
  - from Norwegian Social Science Data Services
  - a viable tool
  - a software system for publishing data on the Web
  - was already being used
    - Europe
    - UK
    - US
    - Canada

- In Ontario, already in use
  - Carleton, Guelph, Ottawa, Queen’s, Toronto, Windsor
  - Statistics Canada
What is Nesstar?

- DDI compliant
- Supports multilingual metadata and multilingual user interface
- Supports both microdata and aggregate data
- Allows searching by keyword for surveys and survey questions
- Does basic data exploration and analysis on the web
- Can download full datasets or subsets in popular formats
- Can export tables and charts
The <odesi> planning group applied for and received funding from the BPS Supply Chain Secretariat, Ontario Ministry of Finance.

In 2007
- a team was struck
- a Project Manager was hired
- 2 part-time Developers were hired
- 2 co-op students were hired

Lead institutions were Carleton and Guelph, with in-kind assistance from Queen’s.

Ottawa and Laurentian joined the team to develop French screens and metadata.
The Process

- **First step**
  - to develop a Canadian ‘Best Practices’ document for cataloguing data files using DDI

- **Next step**
  - students ‘marked up’ (catalogued) surveys which were initially loaded onto a test server at Guelph
    - StatsCan PUMFs and aggregate files
    - Canadian Gallup Polls

- **In 2008**
  - hardware and software were purchased through Scholars Portal
  - the team at Scholars Portal worked with the <odesi> team to establish a data server and to load data files
Next steps …

- `<odesi>` was launched by Scholars Portal March 4, 2008

- Negotiated access to the IPSOS Reid public opinion polls at LISPOP (Wilfrid Laurier) for use in `<odesi>`
  - `<odesi>` funded students marked-up the polls

- Along the way
  - the DLI team offered help and suggestions
As well ...

- Engaged in many promotional activities
  - created <odesi> brands
  - produced bilingual brochures and posters
  - presentations at OLA, CLA, IFLA, IASSIST......
- Extended the contract with the <odesi> developers for a second year
- Hired a programmer at Scholars Portal to further customize and improve <odesi>
- Completed the ‘Best Practices Document’ for DDI data coding in English
  - had it translated into French
- Looked for new datasets
  - Province of Ontario Data
  - ICPSR
  - Voice of the People

- Created a suite of tutorials and other training materials in both English and French

- Investigated providing access to <odesi> to the wider education sector
  - Colleges
  - Universities across the country
- Regular meetings to make sure everything was on track
  - Management
  - Operations
  - Partners
  - Developers
- Additional contributions to <odesi>
  - Supervision of the work process at various universities
  - Carleton
    - In-kind through extra student power
  - Queen’s
    - Full-time staff member marking-up part-time
“Initiative” was finished

<odesi> became a permanent service of Scholars Portal

Scholars Portal hired a Data and Geospatial Librarian in the summer of 2009 to manage <odesi>
  - Leanne Hindmarch

Term for the Project Manager was over in the fall of 2009
MarkIt! Program

- Scholars Portal makes monies available to Ontario universities to fund students marking-up files to go into <odesi>
  - Carleton
  - Guelph
  - Laurentian
  - McMaster
  - Scholars Portal
  - U of Ottawa
Features: data holdings

- **Bilingual DLI Microdata includes**
  - Census
  - Special Surveys
  - General Social Surveys (GSS)
  - Health

- **Bilingual DLI Aggregate Data includes**
  - Justice
  - Health
  - Census
  - Education

- **Other Microdata includes**
  - ISSPs – Canadian data
  - Canadian Addiction Survey
  - Surveillance Project
  - Canadian Millennium Scholarship Foundation
Data holdings (cont’d)

- Public Opinion Polls include
  - Gallup Canada
  - IPSOS Reid
  - Listening to Canadians

- Links to metadata include
  - ICPSR
  - CORA (Queen’s)

- As of the end of April 2011
  - Records from over 11,433 surveys

- New surveys added regularly

J.Fry, S.Mowers, DLI Boot Camp, May 31 2011
Other features

- Choice of interface language (English/French)
- 2 mirror servers
  - load balancing leading to quick response time
  - Security
- An excellent, bilingual search engine developed by Scholars Portal based on Marklogic
- RSS feed
- A centralised-managed email address that responds to queries and problems quickly
Features (cont’d)

- **Search**
  - Refine by
    - Date (single or range)
    - Your collection of choice
  - Combine multiple search concepts
    - Anywhere
    - Variable/Category Label
    - Survey Title
    - Keywords
    - Abstract
    - Series

- **Browse by**
  - Category
  - Keyword
  - Survey series
  - Title
Behind the scenes ...

- Excellent tech services crew

- OOG – the <odesi> Operational group
  - meets regularly, usually monthly
  - includes members from all Ontario universities

- Mark-up Group
  - meets as needed
  - addresses issues, concerns, …
Future Directions

- Becoming a Trusted Digital Repository
  - Presently in the process of being examined

- Subscriber access for institutions outside OCUL
  - Colleges
  - Non-profit organizations
  - Universities outside of Ontario
  - Contact Leanne Hindmarch for details
    l.hindmarch@utoronto.ca
Here it is!

Let’s check it out!

www.odesi.ca
Contact

For more information:

Jane Fry
Jane_Fry@carleton.ca

Susan Mowers
smowers@uottawa.ca