Building on the first Government Information Day held November 1, 2013 at the University of Toronto, this second edition offered 17 presentations grouped under 4 sessions: Digitization, preservation and access; Web harvesting; Electronic government information and social media; and Open government in Canada.

Welcome: Leslie Weir, University Librarian, University of Ottawa
Weir pointed to the importance of government information to society and the role of librarians as custodians and mediators of this information. This role presents many challenges and the theme of the conference was how they might be met through collaboration, both at the institutional level and between individual librarians. Finally, Weir emphasized how this forum is important to keep abreast of the landscape – past and future – as well as emerging and ongoing.

Session 1: Digitization, preservation and access

Digitizing Ontario government documents: The story so far
Simone O’Byrne [Chair, Ontario Government Libraries Council Working Group on Ontario Government Publications] Loren Fantin [Director, OurDigitalWorld]

In spring 2013, Our Digital World worked with OCUL’s Ontario Digitization Initiative and an iSchool practicum student to survey 77 Government Documents Librarians on digitization projects of Ontario Government publications. The intent was to make the information available via a wiki. Twenty two responses were received, but not all questions on each questionnaire were answered. The questionnaire sought to capture organizational information; how the project was being undertaken by way of partners, workflow, and funding; the status of the project, whether proposed, on-going or completed and how access was being provided to the results. Preservation was identified as an impetus for digitization in more than half of the cases. Lessons learned from some digitization projects: 1) high quality metadata is very important and 2) access to documents is proving to complement reference services, which have been seriously cut back in some departments. If no digitization project was underway, the reasons given were: staffing, lack of funding, low organizational priority, lack of technical skills, technical infrastructure and new organizational structure. In 2014, the survey was redesigned and focused on collecting information to be the basis for a registry with search and display functionality. The questionnaire has not yet been sent out. Future partners may be the Hathi Trust, University of Ottawa, and the University of Toronto. Specifically with reference to the Government of Ontario, there is an interest in digitization for reasons of space, budget, and continuing access. However, the obstacles are limited resources, financial and human, a demanding procurement process, and inflexible/inconsistent IT policies. In addition, not all departments/agencies have a library. There is a concern over the number of publications that will be lost with the migration to Ontario.ca. Here the emphasis will be on what is in demand and current.

Who’s digitizing what?
Catherine McGoveran [Government Information Librarian, University of Ottawa]

General interest in a registry of digitization projects was expressed at the 2013 Government Information Day meeting. Such a registry, would enable libraries to find collaborators, identify complementary collections, avoid the hazard of duplication of effort, and would facilitate sharing. Requirements for the registry are that it is bilingual, easy to use in terms of inputting information, and allows data export. The U.S. FDLP was looked at
as a model. Institutions will need to assume responsibility for the currency and quality of the information submitted. So far there have been 12 project level submissions. A beginning has been made and co-operation is sought for publicizing the initiative.

Planning for a registry of digitized Canadian government documents
Margaret Wall [Communications Librarian, University of Toronto]

The purpose of this proposed registry is to provide users with a single access point to publicly accessible digitized Canadian federal, provincial, and municipal government documents and to enable institutions considering digitization projects to determine what work others are contemplating, have in process, or have completed. All institutions are welcome to participate. The intention is to migrate the information collected from this project and others into a single database with search and browsing capability. Various metadata models are currently being evaluated. Born digital publications are beyond the scope of this proposed registry.

The 30-year rule has been gone for just over 30 years now: do we miss it?
Fabien Lengellé [Director General, Services Branch, Library and Archives Canada]

Under the new Access to Information and Privacy (ATI&P) legislative framework, LAC is responsible for the retention of classified government records. Currently, there are 460 km records at LAC, 35%-40% of which are still kept secret. Under the ATI&P, many exemptions do not have time limits and LAC must pro-actively declassify them through a block review process, involving an “injury test” which requires an assessment of the risk of release in terms of the existence of real and imminent harm to the Government of Canada or the Canadian population. According to Lengellé, the open government initiative is eroding the injury test. The process of declassification is slowed by various issues such as the rate of creation of government records, the variety of formats, the inherent limits to block review process (done by sampling files) and the fact that the description of records is at the series level only.

Cuttings or compost? Working with the weeding outfall of LAC and GoC
Amanda Wakaruk [Government Information Librarian, University of Alberta]

The University of Alberta library received 393 boxes of Government of Alberta publications (2nd copies) discarded by LAC and embarked on a major digitization project with the cooperation of the Legislative Library of Alberta and the Internet Archive. This project requires a lot of time, effort, and dollars. Although the project has not yet been promoted, the usage statistics are high. The library is keeping one print copy.

Another ongoing project, involves a systematic study of the removal of documents from the websites of three major departments: Industry Canada; Citizenship and Immigration and Health Canada between 2005 and 2014. There are currently no tools to help identify web content which has been removed. The study is very labor intensive and involves a case by case assessment “with your own eyes”. Work has begun on checking the fate of databases which appeared in Supplementary Checklists 05-03 and 08-01.

Digitization of the Statistics Canada Library's Historical Collection
Jennifer Pagnotta [Chief, Statistics Canada Library]

The Strategic Review 2009/10 recommended digitization of DBS/Statistics Canada official print publications for reasons of preservation and improved access. Digitization began with the census and, in October 2011, the process continued with the 90 series and will move to the 10 series. Records in the Library’s catalogue indicate
whether a publication is “done” and can be requested. It is estimated that work on the 10 and 20 series, missed items, donations, and oversized materials with inserts will be completed by 2015. Complicating factors are the poor condition of some documents, workload from numerous donations, incomplete catalogue records for older items, funding, workforce adjustment, and a change in library systems. The file’s size means that the content cannot be hosted on the website. DLI might provide an interim solution to dissemination.

**Historical Debates of the Parliament of Canada Portal: Increasing access to Parliament’s documentary history and heritage**
Janet Bennet [Library of Parliament]

In conjunction with Canadiana.org, the Library of Parliament has developed a portal which provides public access to all the debates of the Parliament of Canada from 1867 to 1996, the point at which the debates of both Houses are available at the website. The project had many complex challenges such as the integration of pre-1875 debates and recent translations of debates produced only in English, as well as varying levels of quality in print material and metadata. The Journals of both Houses will likely be digitized next.

**Digitization, preservation, access: Canadiana.org and trustworthy digital repository (TDR) services**
William Wueppelmann [Chief Information Officer, Canadiana.org]

Canadiana.org is seeking TDR designation from OCLC and CRL, a process which began in 2012. There is no specific definition of TDR, but there is a checklist of characteristics including organizational infrastructure, digital object management, technology infrastructure, and security. The weighting for assessment purposes is 70% policy and governance, and 30% technical. The decision has been made to certify existing capability and to add new capabilities later. The timeline at the time of the presentation was:
- **May 2012 - February 2014** Self audit
- **March 2014 – June 2014** Submission evaluation
- **June 2014** Two day on site audit
- **June – October 2014** Follow-up and certification

Canadiana.org is using open source software, there are 5 copies of the files in 3 locations. The files are verified once a week. There are automatic alerts for missing/corrupt files and manual intervention to fix errors which have been detected. There are 150,000 archival packages and 50 terra bytes of data per copy. In terms of capabilities and services, Canadiana.org can provide digitization (4 file formats supported at the moment JPEG, JPEG 2000, TIFF and PDF), metadata creation, and conversion. This is all done on a cost recovery basis. Future developments involve metadata enhancement (possibly through crowd sourcing) and more nodes across the country.

**Session 2: Web harvesting**

**Web archiving: Getting started with Internet Archive's Archive-IT service**
Andrea Mills [Collection Specialist, Internet Archive]

Begun in 2006, Archive-IT is a subscription web archiving service from the Internet Archive which helps organizations to harvest, build and preserve collections of digital content. At present, there are 468,000 Canadian items of which 20,000 are government publications. A subscription includes access, storage, and tools. One can request single or multiple crawls. Social media feeds are among the types of resources that can be captured. Heritrix is the web crawler and Umbra gives the flexibility to access sites with bots. Nutchwax is used for searching. When considering archiving, an institution needs to consider its mission and mandate to
collect. Goals and objectives need to be developed both for the short term (3 years) and for the long term (10 years). This process involves deciding on the vision for the collection – what it should look like and how it is to be managed, maintained and used. Collections may be broad e.g. Canadian government (University of Toronto) or specific e.g. Ebola (University of Manitoba). Those wishing to get started can set up a test account (5 URLs and 1 crawl) and practice. The scope and frequency of crawls will determine the required level of subscription.

Cultivating community engagement through web archiving
Nicholas Worby [Government Information & Statistics Librarian, University of Toronto]

Web archiving electronic government information is a labour intensive process. Sites must be selected and a test planned. After the crawl, there is automated quality assurance, manual quality assurance, and the creation of metadata. The University of Toronto Libraries crawl of over 600 sites (Ontario, City of Toronto, and selected federal sites) produced 2 terra bytes of data for 2 librarians and 4 students to process. Five Faculty of Information Studies students were willing to assist. They were given one hour of training and a manual, and it was believed that each could finish the QA on a single provincial site in three hours. It became clear that more time was needed and that there was a different understanding both of the scope of the project and of how web archiving works. This experience led to ideas for improvement and the development of new strategies. It was learned that students should be given more opportunities related to the archiving process, more training in its operation, increased feedback and mentoring, and clearer guidelines. It now appears preferable to choose thematic collections with a smaller scope, such as the Toronto municipal election. Working on digitization with the Archives of the City of Toronto required the securing of permissions, the defining of responsibilities, training, and support. There was a difference in views on the operating principle behind metadata construction, organizational structure (Archives) versus classes of information (University of Toronto Libraries). Archiving projects will need to have complementary goals, be realistic in scope, and have public appeal.

On the hunt for fugitive government information
Sam-chin Li [Government Information Librarian, University of Toronto]

Fugitive documents are those not collected by the Depository Services Program e.g. those in HTML and those that go unreported to the DSP. There is a lack of bibliographic tools for tracking these publications. There is also a lack of effective monitoring of compliance. Electronic publishing contributes to the low awareness of the need for compliance. Also considered fugitive publications are those from quasi-government agencies, which fall under the Financial Administration Act. These publications should be catalogued so the contents are accessible. They should also be archived, so that the contents are preserved. Fugitive documents should be reported to the DSP and added to the repository. The Fugitive Document Working Group, which is a sub-group of the Canadian Government Information – Private LOCKSS Network conducted a survey to find out whether institutions were collecting Canadian government fugitive material and, if so, how it was being done. The survey found that to varying degrees, academic libraries were harvesting these documents using a variety of means for capture. The University of Toronto has archived the news.gc.ca site, the Office of the Information Commissioner’s completed access to information requests, and the Find Statistics and Data by Department page. The creation of a “Fugitive Documents” section in the Internet Archive could help to fill the current gap.

Session 3: Electronic government information and social media

Keeping current, staying relevant
Maureen Martyn [Chief, Current Awareness and Strategic Analysis Section, Library of Parliament]
This year the Research Service marks its fiftieth anniversary and one of its most pressing challenges is to remain responsive and current. The service is exploring many solutions. For example, there will be a pre-release of Legislative Summaries (no editing) on the LEGISinfo site. Users will also begin to see Legislative Summaries for private members’ bills after second reading. Hill Notes are brief analyses of potentially significant issues and have been available to the public since 2010. There is a plan to create a Hill Notes Blog, which would provide key messages at a glance. There is an archive of Hill Notes (older than 90 days) under each category of Current Publications. Other projects discussed are the creation of infographics and embedding current and emerging issues in the Parliamentary website.

Public library micro-blogging and community engagement: Results of a national survey
Mary Cavanagh [Assistant Professor, School of Information Studies, University of Ottawa]

A public library may be considered a government agency because of the nature of its funding. Social media offers libraries frameworks for engagement and participation. A survey was undertaken of 71 larger public libraries in Canada about their use of Twitter and how they perceived its challenges and benefits for engaging with their communities. The benefits appeared to be the relationship/connection that was forged and how Twitter could be used for promotional purposes. The challenges were: the amount of organization and manpower required, time required, effectiveness/efficiency, audience management, determining content, and avoiding controversy. More information can be found at Social-biblio.ca, which tracks and preserves the Twitter activity of Canadian public library systems and those who engage with them.

E-informing the public: Information access and e-government
Luanne Freund [Associate Professor and Chair of the iSchool PhD program, University of British Columbia]

Three separate studies were conducted focusing on producers, the public, and librarians. The first was a small scale project of interviewing government employees. Their goal emerged as being one of information dissemination. Also, they saw engaging and educating the public as being in keeping with an increasing awareness of government accountability. Themes that emerged were: the size and complexity of government information; the difficulty of making the environment simple, consistent and intuitive; designing for the customer with layered content that would accommodate the specialized stakeholder; designing complex sites for browsing; and the limited value of metadata for keywords and topics.

The next project involved thirty three interviewees who were asked why and how they searched for government information. Most of the searching was for short term personal reasons or part of a larger process. Many considered looking for government information difficult. When searching for their own information, satisfaction with results declined, with many pages considered to be unhelpful. Survey subjects had difficulty discriminating between non-government sites that looked like government ones and genuine government sites. There was a low level of search literacy.

The last stage was a survey of 113 librarians. They emerged as mindful of the fact that accessible did not mean more easily found and that patrons might need more help. Issues raised were training, the threat to preservation, findability (not an effective strategy to rely on Google), the digital divide (need for information literacy skills), and the changing role of librarians.

Session 4: Open government in Canada

Open government 2.0: Learning from the past and moving forward
Mary Francoli [Associate Professor of Communication Studies, Carleton University]
This year the Government of Canada, a member of the Open Government Partnership, is working to develop its second action plan on open government. The new action plan comes at the end of an independent assessment prepared by Dr. Francoli under the auspices of the OGP’s Independent Reporting Mechanism. The cycle involves development, implementation, and government self-assessment. With respect to assessing implementation, the questions are: what were the commitments the government made and are they complete and do they matter? Canada’s progress report 2012-2013 finds that the plan needs to be more ambitious. There is a need for goals that require real change. There also needs to be a better foundation for open government in Canada that involves Access to Information legislation, the Office of the Information Commissioner, improved information management and disclosure policies, and more stability. Challenges are: the organizational culture within government and the public service, the multi-jurisdictional make-up of the country, the need to develop a “Canadian” solution, lack of awareness of the Open Government Partnership, and the strength of civil society.

Canada's Second Action Plan on Open Government
Mark Levene [Acting Director, Policy and Liaison for Open Government, Treasury Board Secretariat]

The government is anxious to raise general awareness of the Open Government Partnership and to engage in more constructive collaboration. Information is to be treated as a valuable resource. Evidence-based policy making, the health of the economy, culture, and national identity depend on how information is treated. Canada’s 2nd Action Plan on Open Government is looking at specific language to define the government’s twelve commitments, more ambitious commitments, in three areas: open data, open information, and open dialogue. There were three phases of consultation: how to run the consultation; determining what an ideal plan would look like, and collaborating with other levels of government and the private sector. Concerns are: data availability and quality, data literacy, citizen engagement, promoting transparency, and supporting collaboration. An Open Data Working Group is looking at how success should be measured, long term preservation issues, and metadata standards (98 open data principles).

The conditions and challenges for sustainable open data ecosystems
Jean-Noé Landry [Director of Strategic Initiatives and Outreach, Open North]

Landry introduced two Open North websites: Citizen Budget and OpenHouse Nova Scotia designed to promote collaborative governance. Also mentioned was QuebecOuvert and its municipal off shoots. In Landry’s view, the time has come to move beyond hackathons to issue based problem solving. Champions must be found to support the idea that open data has a value which can be unlocked by businesses and non-profits in the interests of social innovation and collaborative governance.

Summaries of presentations provided by Frances Montgomery and Sylvie Lafortune, both librarians at the Carleton University Library.