3-28-2017

The Canadian Linked Data Summit: Developing Canada's Linked Data Future through Cooperative Alliances

Jennifer J. Browning
University of Toronto Libraries, jennifer.browning@utoronto.ca

Robin Elizabeth Desmeules
McGill University Library, robin.desmeules@mcgill.ca

Sharon Farnel
University of Alberta Libraries, sharon.farnel@ualberta.ca

Andrew Senior
McGill University Library, andrew.senior@mcgill.ca

The Canadian Linked Data Summit: Developing Canada's Linked Data Future through Cooperative Alliances

Abstract

From October 24 to 26, 2016, the Canadian Linked Data Initiative (CLDI) hosted the Canadian Linked Data Summit in Montreal, Quebec with the goal to increase awareness and nurture collaboration for linked data production in Canada. The Summit was inspired by CLDI's investment in developing and sustaining a cooperative plan for Canadian linked data development for libraries, archives, museums, and other cultural institutions across the country. CLDI, comprising of Canada's five top research libraries, the University of Toronto, McGill University, Université de Montréal, University of Alberta, and the University of British Columbia, and partners at Library and Archives Canada, Bibliothèque et Archives nationales du Québec, and Canadiana.org, organized the CLDI Summit to allow library staff specializing in cataloguing and technology from institutions across Canada to become better equipped for opening our library metadata to the global Web through the production of linked data. Gathering together linked data experts from North America and Europe, librarians from academic, government and special libraries, as well as graduate students from Canadian Library and Information Science schools, the CLDI Summit provided a forum for recognizing the importance of linked data for libraries, sharing expertise and resources, and working collaboratively between units and institutions across the country. Consisting of presentations and panel discussions, hands-on workshops, and a stakeholders planning meeting, the 3-day CLDI Summit helped to ignite and sustain real strategies for how to move forward with linked data knowledge and production in Canada through leadership, collaboration and communication.

From the Field
The Canadian Linked Data Summit: Developing Canada's Linked Data Future through Cooperative Alliances

Jennifer J. Browning (jennifer.browning@utoronto.ca)
Metadata Librarian for Electronic Resources, University of Toronto Libraries

Robin Elizabeth Desmeules (robin.desmeules@mcgill.ca)
Cataloguing Librarian, McGill University Library

Sharon Farnel (sharon.farnel@ualberta.ca)
Metadata Coordinator, University of Alberta Libraries

Andrew Senior (andrew.senior@mcgill.ca)
Assistant Librarian & Coordinator of eResources and Serials, McGill University Library

Background

With increasing momentum, libraries and cultural institutions around the world are focusing their attention and resources on developing new standards and practices for implementing linked data. The eventual replacement of the MARC standard for library bibliographic description provides the opportunity for bibliographic and archival metadata to become exposed beyond the confines of closed databases and library systems. By leveraging new technologies and alternative metadata schemas, linked data allows us to create interconnected digital spaces that facilitate new avenues for user discovery and means of accessing our library resources. In Canada, despite the awareness that linked data is likely to be the next phase of resource description, and though there are linked data projects currently underway, there is a realization that we are falling behind our counterparts in the United States and Europe in terms of developing coordinated strategies for linked data preparedness on a national scale.

Motivated by the realization that Canada may be behind the curve in linked data development, the Canadian Linked Data Initiative (CLDI) was collectively constituted by its members with the aim to develop a communal approach for linked data development in Canada. By joining together individual library departments and cultural institutions across Canada, CLDI seeks to nurture and sustain a culture of open metadata creation and the application of machine-linkable technologies through collaboration. Focusing on resource sharing, education and awareness, and project development, CLDI strives to ensure that Canadian libraries and cultural institutions are equipped to not only keep up with the inevitable changes to bibliographic description and digital technologies, but that we are major contributors to the global linked data movement.

CLDI had its inception in the spring of 2015. Periodically, the collections and cataloguing directors from the top five research libraries in Canada – University of Toronto, McGill University, Université de Montréal, University of Alberta and University of British Columbia –meet via teleconference and in person to discuss common issues within their libraries' technical services departments, such as technology, collections budgets, and staff training. Through these dis-

cussions, it became evident that a major challenge affecting these libraries, and many libraries across Canada, is the inevitable shift from MARC and AACR2-based cataloguing toward new metadata standards that interplay more intelligently with digital environments. While there has been excitement about new cataloguing standards like BIBFRAME 2.0 and Resource Description and Access (RDA), employing these standards with linked data presents libraries with many challenges, including staff training, information technology support, and changing workflows. In comparison to the international community, it became clear that in many respects, Canada is behind in its capacity to produce linked data content within our libraries' discovery tools, catalogues, and digital collections. By the fall of 2016, the need for a Canadian plan for linked data was met with mounting urgency, thus inspiring the establishment the Canadian Linked Data Initiative (CLDI) to develop a coordinated linked data strategy for libraries in Canada. Soon after its inception, Library and Archives Canada, Bibliothèque et Archives nationales du Québec, and Canadiana.org joined CLDI, establishing the group as a bilingual national body that represents Canada with broad institutional participation. Aligned by this formal partnership, the members of CLDI have set out a mission to collaboratively plan, facilitate, and obtain funding for linked data projects in Canada.

Goals and Objectives of CLDI

It is fitting that CLDI leverages the collaboration of its partners in order to implement linked data projects in Canada, as in principle, linked data itself depends on open data and resource sharing to propagate interconnected digital spaces. To improve the discoverability of library resources and digital collections, successful linked data projects rely on the establishment of rich networks based on asserted relationships between works, authors, and other entities. This cannot be done without the broad participation

of libraries to expose their metadata to the Web by linking controlled vocabularies and unique resource identifiers (URIs).[1] By working across departments, institutions, and languages, CLDI aims to develop strategies for ensuring that this rich data is shared by Canadian libraries on the open Web.

To being this important work, CLDI discussed the various factors impeding the production of linked data in Canada and the scope of these obstacles. They then prospected resources to support this initiative in the form of institutional investments and grants. This initial information gathering and program development was directed by the Steering and Planning Working Group, comprised of the Associate University Librarian or equivalent role for each CLDI member institution. Additional working groups were then established to support specific initiatives and requirements such as funding, education and training, and project planning. Functioning under the general direction of the Steering and Planning Working Group, these working groups included the Digital Projects Working Group, IT Working Group, Identifiers Working Group, BIBFRAME Working Group, and the Education and Training Working Group. Once these groups were formed, they were tasked with establishing their own mandates and arranging meeting agendas based on initial goals they had set for 2016. With the aim to develop capacities in technical services departments across Canada, CLDI members are able to work collaboratively within these groups, despite encountering particular challenges such as increased workloads and geographic differences, with the shared goal to increase linked data awareness and alleviate the real or perceived impediments to linked data production in Canada.

The Canadian Linked Data Summit

Realizing that a linked data strategy necessarily involves the collaboration of not only CLDI

members, but the participation of the wider Canadian library community, one of the first priorities undertaken by the group was to provide a forum where professionals and linked data experts could gather in person. Held on October 24 to 26, 2016 in Montreal, Quebec, the Canadian Linked Data Summit was organized by CLDI to allow scholars and practitioners from libraries and cultural institutions across Canada to demonstrate their linked data projects, share their expertise, and devise strategies for sustaining linked data development in Canada. By asking librarians, library staff, library school graduate students, and other parties interested in linked data to meet in person to discuss methods of moving forward, as well as inviting international linked data experts to share their expertise and speak about their projects, the Summit took a great step toward abating the difficulties of working across units, across institutions, and across languages in Canada. Agreeing that the development of a linked data strategy requires the broad participation of libraries across Canada, CLDI members felt that hosting the Summit was an essential step towards developing and sustaining linked data creation in Canada.

Collaborative Planning

Planning and fulfilling the goals of the Canadian Linked Data Summit could not have been achieved without a great deal of coordination and communication within and between CLDI working groups to decide on the Summit's content and to make the various logistical arrangements required to host the event. Even without formal advertising, the Summit generated a great deal of interest from library staff across Canada, and its planners were encouraged to see that registration numbers had reached over 100 participants in early October. Certainly, one of the most important factors in planning such an event was to obtain funding to pay for the various costs of the Summit. In order to ensure that the event was accessible and inclusive to the

broader Canadian library community, it was important that individuals could attend the Summit without incurring significant expense. Thanks to the resources obtained by the Grants Working Group, CLDI received financial support for the Summit through the generous backing of CLDI partners, as well as a Connection grant from the Social Sciences and Humanities Research Council (SSHRC). Due to this funding, CLDI removed registration costs for the Summit, so that travel and accommodation costs were the only expenses incurred by participants and their home institutions. Additionally, since linked data development in Canada requires the involvement of not only the library community, but also new professionals entering the field, CLDI felt it was important to sponsor selected graduate students from Canadian Library and Information Science schools to attend the Summit as a way of encouraging new librarians to become part of Canada's linked data future. The Grants Working Group and other CLDI members were instrumental in not only ensuring that there was enough money to fund the event, but that the resources were appropriately allocated between various expenses, such as travel and accommodation for speakers, room and equipment rental, a catered reception, meals for attendees, and translation costs between French and English. Once bolstered with finances to host Summit, CLDI working groups forged ahead with organizing the agenda and inviting linked data experts from across Canada, the United States, and Europe to participate in the event.

Based on CLDI's goals of raising awareness, sharing of expertise, and collaboratively planning linked data strategies, the Canadian Linked Data Summit comprised three days of presentations, panel discussions, workshops, and meetings. It was initially intended that Day 1 and Day 2 would be open to all attendees, with the first day consisting of presentations and panel discussions involving linked data innovators

from across the US, Europe, and Canada, and the second day involving practical hands-on workshops. Day 3 was set as a planning day for CLDI project staff, speakers, and invited guests to work towards drafting a coordinated linked data strategy for Canadian libraries. Held at the Université de Montréal and McGill University in Montreal, Quebec, planning the Summit involved a strong coordinated effort of all working groups and their members. The Steering and Planning Committee held regular conference calls to delegate work and discuss the logistics for planning the Summit. In addition to these regular meetings, an intranet page was created to facilitate brainstorming, and allow for the organization and sharing of documents. While a great deal of the planning was accomplished remotely, librarians working both at Université de Montréal and McGill University worked tirelessly to arrange local particulars in Montréal. Individual and team efforts at both institutions contributed significantly to the overall success of the Summit, from making hotel arrangements and restaurant reservations, to handling with swiftness technical problems and other minor glitches that undoubtedly arise when planning events of this scale.

Day 1 – October 24, 2016: Collaborative Linked Data Projects from Around the World

The first day of the Summit, held at Université de Montréal, was billed as "Collaborative Linked Data Projects from Around the World". This day gave attendees the opportunity to learn more about linked data through discussions and presentations given by linked data experts from the United States, Europe, and Canada. Each presentation was 20 minutes in length, with time provided for questions and discussion. The day began with a welcome and introduction by Caitlin Tillman from the University of Toronto, Julie Cardinal from Université de Montréal, and Joseph Hafner from McGill University, all of whom are members of CLDI Steering and Planning Working Group and were instrumental in

initiating the Canadian Linked Data Initiative. After these introductions, the day proceeded with presentations and demonstrations of linked data projects from Canada and abroad. The morning comprised of presentations from the following speakers:

● Phil Schreur, Stanford University Libraries: "Linked Data for Production: A Multi-Institutional Approach to Technical Services Transformation"

● Etienne Cavalié, Bibliothèque nationale de France: "Catalogs and the Linked Data Challenge"

● Xiaoli Li, University of California Davis Library: "UC Davis Grant: Reinventing Cataloging: Models for the Future of Library Operations"

These presenters then fielded questions from the audience and led a discussion that focused on the importance of building partnerships not only within the library community, but also with vendors and IT departments to ensure that we are equipped with the technological infrastructure needed for developing linked data projects. After lunch break, attendees reconvened to learn about more linked data projects from several more experts in the field:

● Jeff Mixter, OCLC: "Linking Your Data"

● Marielle St-Germain (speaker) and Jean-François Gauvin (co-author): "Données Ouvertes et Web de Données: Enjeux et Perspectives pour BAnQ"

● Rob Warren, Muninn Project/Canadian Writing Research Collaboratory: "Operationalizing Linked Open Data"

● Thomas Hall, ExLibris: "Putting Linked Data at the Service of Libraries: ExLibris – A Partnering Perspective"

● Jean Delahousse, DOREMUS Project: "Doing Reusable Musical Data"

● Julienne Pascoe, Canadiana.org: "Canadiana.org: Linked Data Practices & Collaboration"

● Christina Harlow, Cornell University: "Pragmatic Efforts towards Linked Data Infrastructure"

● Stacy Allison-Cassin, York University: "The Linked Data Cloud, Cultural Heritage, and the Canadian Context"

● MJ Suhonos, Ryerson University: "Linked Data in Canada: Behind the Curve"

Indeed, it was a busy yet inspiring day that generated a great deal of discussion and ideas for how to move forward with linked data production in Canada. The theme of establishing partnerships that emerged from the morning sessions was sustained throughout the afternoon, stressing the need for collaboration across institutions if we are to break away from the silos created by our traditional library systems and MARC bibliographic standards. The examples of linked data projects demonstrated in these sessions provided attendees the opportunity to see how, through experimentation and cooperation, the future of linked data may take shape in Canada.

Day 2 - October 25, 2016: Canadian Linked Data Workshops

The second day of the Canadian Linked Data Summit was designed to facilitate the sharing of expertise and increase practical knowledge for librarians and staff working in technical services departments. When planning the agenda for the day, it was decided that CLDI working groups would each develop a practical workshop with the intention to develop linked data competencies for Summit attendees.

Workshop 1, hosted by the Digital Projects Working Group, allowed attendees to reflect on key aspects of planning and executing a linked data project as they brainstormed potential challenges and opportunities that may arise when

undertaking a linked data project in their home institution. The working group presented a proposal for a shared digital project based on the theme of Canada's 150th anniversary of Confederation in 2017, and outlined the steps they are taking to plan and execute the project. Inviting Summit attendees to collaborate in this project, the working group plans to use graduate student theses from Canadian academic libraries to expose interesting connections between places, time periods, subjects, and authors using linked data, with the hope that linked data may uncover new stories about Canada and its history of academic study. Generating a great deal of interest from libraries outside CLDI membership, this workshop provided the Digital Projects Working Group with new ideas of how to move forward with a project of this national scope.

Workshop 2 was organized by the IT Working Group, who provided attendees with another opportunity for active participation through a brainstorming session. The session encouraged participants to consider the various technology-related questions that must be answered when developing linked data projects, as they often require the involvement of IT departments and/or technology specialists. Participants were provided with several case studies to help guide them through the thought process and were asked to chart what key decisions or tasks need to be made when implementing a linked data project in order to make smart technological decisions. This session demonstrated the importance of collaborating with areas outside of library technical services and the need to consult the expertise of library IT and systems departments when developing linked data projects.

Workshop 3 comprised a presentation from the Identifiers Working Group, who demonstrated a variety of tools that can be used to create linked data-ready metadata through identifier reconciliation, RDF generation and legacy metadata en-

hancement. The applications presented included MARC Next, OpenRefine, RIMMF (RDA in Many Metadata Formats), and Protégé – some of the linked data tools available today. The session aimed to describe both the purpose of these tools, as well as to provide considerations for how they can be potentially applied in technical services workflows. As several of these tools can be challenging to work with, this workshop also underscored the importance of documenting both our successes and our failures when experimenting with linked data tools. In many cases, trial and error is the best way for us to learn how to augment our metadata with linked data. By sharing our use of these tools with the library community, we can begin to see more real examples of linked data in production, as well as encourage the development of tools that are more readily adoptable into our library workflows.

In Workshop 4, the BIBFRAME Working Group explored linked data creation tools for bibliographic description using BIBFRAME. BIBFRAME, now succeeded by BIBFRAME 2.0, is widely seen as a potential replacement for MARC21 bibliographic format due to its development of standards using a linked data model. At the time of the CLDI Summit, there were two main tools that allowed for metadata conversion to BIBFRAME – BIBFRAME Editor and Scribe from Zepheira. The session provided a short demonstration of each tool and a critical analysis of their outputs. The working group discussed advantages and opportunities for both tools, provided examples of metadata in RDF serialization format, and presented a demonstration of the MARC to BIBFRAME Comparison Service. Corroborating the conclusions made by the Identifiers Working Group, this workshop stressed the importance of librarians to advocate for BIBFRAME tool development and to urge library service platform providers and publishers to become more active agents within the linked data supply chain. Again, this is work that must

be done in collaboration with the library community and simply cannot be accomplished if we work in isolation.

Finally, Workshop 5 consisted of a panel forum led by the Education and Training Working Group that included several speakers from Day 1. This session allowed participants to share both existing and anticipated challenges when employing linked data. The ensuing discussion exposed the need for more formal training in linked data both for existing library professionals and for Library and Information Science students. Developing a set of core competencies for technical services librarians and support staff would allow individuals to develop the necessary skills for creating linked data and allow the Canadian library community to establish strategies for working within a post-MARC era. This final workshop highlighted many of the opportunities and challenges that impede linked data production in Canada and how they may be mitigated by working more collaboratively across units, across institutions, and across the country through documentation, resource sharing, and experimentation.

Day 3 – October 26, 2016

On the final day of CLDI Summit, the working groups came together to recap the event, discuss the year's progress, and propose goals for the upcoming year. Once again highlighting the fact that the work of CLDI requires constant coordination by its members and continued communication with libraries across Canada, the day began with a workshop on techniques for collaboration, presented by Salvy Trojman from Gale Cengage Learning. This presentation outlined some of the strategies that Gale uses when accomplishing projects that are often international in scale and that involve many stakeholders, and it provided some suggestions for how CLDI could adapt these strategies when working on their own initiatives. In particular, Trojman described how CLDI could adapt the Agile

software development model, which Gale employs for its own product development with considerable success. The Agile model allows for the progression of products through successive iterative developments cycles that are then added to and revised with new features or improvements based on stakeholder feedback, and Trojman suggested that CLDI could adapt such a model when developing linked data projects and educational tools.[2]

Following this workshop, the members of each working group met to discuss their next steps and goals for the fall of 2016 and 2017 and a roundtable discussion was held. Several new strategies and ideas emerged from these meetings. Recognizing a great deal of convergence of work between the BIBFRAME Working Group and the Identifiers Working Group, the Chairs of these groups proposed to merge them into one working group – the Metadata Working Group – that would cover the various aspects of identifiers, editor tools, vocabularies and data creation. These two chairs also acknowledged that the work of planning and generally overseeing the working groups could be made more manageable, and enjoyable, if the group was led by two co-chairs, which lessens their individual workload and enables collaboration on steering the general direction of the working group. During the roundtable discussion, this co-chair strategy was suggested as being prudent for other working groups to adopt, and several other working groups have since done so.

Another decision resulting from Day 3 meetings was the need to form a User Experience Working Group. As CLDI works towards developing linked data projects in the coming months, there will be a need to test their usability and assess user response to these projects. In the coming months its members will be tasked with testing CLDI projects and initiatives using recognized evidence-based assessment and user experience strategies. Day 3 also allowed for another new working group to convene for the first time – the

Groupe de travail francophone. Due to Canada's bilingual status, it is essential that our national strategy for linked data development has a focused approach for working in and developing documentation in French. At the Day 3 meeting, the Groupe de travail francophone devised its mandate to address the particular challenges posed by adopting linked data technologies in French-speaking communities and outlined some initial strategies and goals for achieving this in the coming months.

In addition to various developments related to the composition of the working groups, discussions on Day 3 underscored the need for each group to collaborate with each other as they move toward achieving their respective initiatives. At times, working groups will need to call on the expertise of other groups when developing and implementing linked data projects. For example, the Education and Training Working group will aid other groups in constructing and disseminating training modules and other educational tools, and the IT Working Group and the User Experience Working Group will work closely with other groups provide technical and assessment support to projects as they are being developed. Collaboration and regular communication between and within working groups will ensure that CLDI meets its goals in a coordinated and organized manner. As the CLDI Summit drew to a close, its members and attendees were energized and inspired by the cooperative spirit that characterized each day of the event.

Lessons Learned & Next Steps

Overall, the Summit was a success for CLDI as it achieved the goal of bringing together the Canadian library community in order to raise awareness and establish strategies for linked data development in Canada. But as with any undertaking of this extent, the positive outcomes resulting from its successful planning and the en-

couraging feedback from attendees were accompanied by some important lessons learned that will help to shape the future of CLDI. While CLDI felt that the Summit was necessary to ignite a broader national interest in linked data, the event helped CLDI members realize the need to become more inclusive in terms of its membership, resource sharing, and communication. To work towards remedying a sense of exclusivity felt by some attendees, the intranet page where CLDI materials are shared between group members has been opened up to the general public so that anyone can consult its resources and documentation as it becomes ready. Additionally, the discussions on Day 3 were opened up to anyone who was interested in attending, and as a result the working groups welcomed some new members from institutions who are not formally part of CLDI. Indeed, CLDI recognizes that linked data development in Canada must include a multitude of voices from a range of perspectives. CLDI will work in the coming months to further engage the library community for further collaboration and resource sharing.

Throughout 2017, CLDI aims to sustain the momentum gained from the Summit as it gets to work developing several linked data projects, such as the Canadian theses project described above, in addition to sharing documentation and training materials with the wider community. In early 2017, each working group will set out goals for the year that involve new linked data projects and initiatives and they will continue to consult with each other when particular skills or expertise are required.

The challenges of working between institutions across Canada will continue to persist, but the adoption of a co-chair composition to the working groups aims to help mitigate some of the logistical issues of working collaboratively on a national scale. Different approaches to holding meetings, such as experimenting with various conference call and video chat technologies, are going to be explored throughout the year, and group members will aim to meet in person when possible, such as at the ALA Midwinter and Annual Conferences in 2017.

Conclusion

In just over one year, the Canadian Linked Data Initiative has achieved a great deal in igniting and sustaining a linked data movement in Canada. Coupled with the successes of the Linked Open Data for Libraries, Archives, and Museums (LODLAM) Toronto Workshop and Summit held in May 2016, the 3-day Canadian Linked Data Summit contributed significantly to linked data awareness, skill development, and project advancement in libraries and cultural institutions across the country. The CLDI Summit was an integral step in working towards CLDI's objective to raise awareness and educate the library community about the importance of linked data, as well as to find ways of mitigating the difficulties of working across disparate institutions, units, and geographic locations in order to sustain strategies for linked data production in Canada. This endeavor, performed on a national scale with a bilingual focus, cannot be achieved without a shared passion for improving access to our library's resources and digital collections. Continuing with a broad, coordinated effort, Canada will ensure its role as a key player in the future of bibliographic description though linked data development.

Authors' Notes:

All presentation slides and additional resources from the Canadian Linked Summit are archived at: https://www.mcgill.ca/clds/.

The Canadian Linked Data Initiative membership, working group documentation, and other resources can be found at: https://connect.library.utoronto.ca/x/OoBKAQ.

1 Virginia Schilling, "Introduction and Review of Linked Data for the Library Community, 2003-2011," *Transforming Library Metadata into Linked Library Data* (Chicago: ALCTS, 2012). http://www.ala.org/alcts/re-sources/org/cat/research/linked-data

2 Salvy Trojman, "Collaboration and Problem Solving in an Agile Environment" (presented at the CLDI Summit, Montreal, October 26, 2016).