Open Data Day Hackathon 2014 at Edmonton Public Library

Alex Carruthers  
Digital Public Spaces Librarian  
Edmonton Public Library  
acarruthers@epl.ca

Abstract

Edmonton Public Library (EPL) hosted its first hackathon for International Open Data Day 2014. International Open Data Day promotes open data policies in local, regional, and national governments worldwide, in the spirit of transparency and civic innovation. The open data movement, like public libraries, values access to information and civic engagement, and it offers opportunities for public libraries to improve their efficiency, transparency, and programming. Celebrating the event provided the Library with the additional benefit of strengthening our relationship to local government. This case study provides a practical introduction to hosting an open data hackathon as a first step to engaging the open data movement. Two follow-up surveys, one immediately after the hackathon and another five months later, were used to assess the event and determine how the Library could better support the open data community in the future. The majority of hackathon participants labelled themselves beginner programmers, were not regular library users, and appreciated the opportunity to meet city employees and other hackers who shared their interests. The Library was encouraged to increase our output of open data and to host more hackathons. Results also suggested room for improvement in the areas of developing a more formal structure to the event, connecting participants with similar interests, and providing long term support for app development. By hosting a hackathon for International Open Data Day, EPL gained both the information and the relationships necessary to release meaningful datasets and put itself in an excellent position to understand and respond to the interests and needs of the open data community.

Keywords

open data; hackathon; public libraries

Introduction

The open data movement has a number of positive implications for public libraries. Were a library to collect and analyze its internal data and integrate it with publicly available data, it could improve the efficiency of workflows and provide evidence-based support for program development. Sharing library data such as in-branch technology
usage, anonymized circulation statistics, and catalogue metadata improves the organization’s transparency and can provide citizens with insight into the value of the library. Open data can also form the focal point of engaging library programming. Offering programming around open data is one way for public libraries to be responsive to the new kinds of literacies and information users emerging in the 21st century. Edmonton Public Library (EPL) took advantage of the annual International Open Data Day as an opportunity to embrace the open data movement and start exploring its potential by hosting an Open Data Day Hackathon.

Inspired by the open data leadership of the City of Edmonton (CoE), becoming involved in the open data community and supporting data literacy are initiatives in EPL’s 2014-2016 Business Plan. The topic of open data has also become increasingly prevalent in professional discourse. For example: a study of 2012 LIS graduates proclaimed the emergence of the “Databrarian” (Maatta); at the start of 2014, open data and big data were listed as top tech trends at the ALA Midwinter Meeting (SinhaRoy); and Greenwalt encouraged public libraries to engage with the trend of open data in a recent article in Public Libraries Online. At EPL, new efforts are guided by our over-arching Community-Led Service Philosophy, which is a framework for building relationships with our community members and developing programs and services based on the needs they express (Edmonton Public Library). Hosting a hackathon was an effort to achieve the organizational goals of participating in the open data community, building relationships within it, supporting data literacy, and understanding the community’s needs.

What is open data?

OpenDefinition.org, a project of the international, non-profit open data advocate Open Knowledge Foundation, explains that "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." Broadly speaking, this describes the accessibility of any information, and the Open Knowledge Foundation encourages the adoption of open licenses by government, businesses, and non-profits. The open data movement has done the most advocacy and made the most impact thus far in the realm of government data. International Open Data Day is a prime example of the movement’s focus on government data: the event website declares that the day’s purpose is to “show support for and encourage the adoption [of] open data policies by the world’s local, regional, and national governments.”

International Open Data Day started in 2010 through a partnership between Canadian and Brazilian open data advocates. Since then, the annual event has grown exponentially, with independently organized events on almost every continent. In 2014 the Open Data Day Wiki listed 194 events held around the world in 47 different countries, including 9 events in Canada (“2014/City Events”). The scope and activities of each event depend on the context. Events in 2014 included public talks and various adaptations of the basic hackathon premise, which responded to local issues and accommodated various skill levels.
A hackathon is a collaborative computer programming event. They do not necessarily involve open government data, as Shujah’s case study of the Steacie Library Dungeon Hackfest demonstrates. They do tend to follow the structure Shujah lays out: “The first half of the event is mainly about forming groups, discussing hack ideas, and getting started on coding. The second half is spent continuing with coding, moving the project forward, and solving issues collectively to create a final product” (2). Hackathon organizers can decide in advance what the general theme of the event will be and can even propose specific problems they hope will be collectively solved by the end of the event. A hackathon that utilizes open data is an example of civic hacking, which is the name given to programming efforts that aim to improve the function of government and the experiences of citizens. What makes hackathons a key tool of the open data movement is their common spirit of sharing and collaboration, which also resonates with the values and goals of public libraries.

**Case Study: Open Data Day 2014**

**Planning**

Documentation on recent Canadian open data hackathons provided essential information which inspired the direction of the event’s development. Open data advocates McArthur, Lainchbury, and Horn have published the “Open Data Hackathon How to Guide,” which provides a thorough and detailed guide to planning and hosting a hackathon. A blog post by internationally recognized open data advocate David Eaves about the City of Vancouver’s 2013 International Open Data Day Hackathon also provided useful insights. Eaves describes the process of “open data-ing” or “speed data-ing”: a play on the concept of speed dating in which city data specialists had quick conversations with local data enthusiasts before they rotated on to a new conversation partner. Eaves writes:

> Open Dataing is about getting everyone engaged. Helping public servants see what citizens are interested in and how they can see technology working for them, it’s also about getting participants to learn about what is available, what’s possible, and what are some of the real constraints faced by city staff.

Feedback on the hackathon published by the City of Vancouver revealed that open data-ing was one of the most popular aspects of the day. Preferring the term speed data-ing, we decided that including this activity in our hackathon would be a compelling way to draw data enthusiasts to our event and an ideal opportunity for EPL to work with the CoE.

**Partnership**

We reached out to the CoE IT department in early November 2013 to propose that we collaborate on the event. Their response was very positive, particularly as community outreach was one of their departmental goals. EPL’s planning team comprised the manager of digital literacy initiatives and web services, the chair of the adult services team, and the digital public spaces intern librarian. From the CoE IT department, the
strategic coordinator, business intelligence team lead, and a communications representative were at the table. We had three months to plan the hackathon. EPL put together a formal proposal that outlined the program for the day (see fig. 1), the responsibilities of each partner, the resources we’d need, the expected costs, and the risks of the project. The successfully received proposal was a useful road map for the rest of the project.

EPL took responsibility for hosting the hackathon in our new Makerspace, an adaptable creative space with sophisticated tools for engaging with digital media. Hosting responsibilities included managing the registration procedure as well as providing coffee and pizza, an MC, a “best idea” prize, and door prizes. We also developed a small promotional campaign and established an assessment strategy. The CoE was responsible for organizing city business unit representatives to attend the speed data-ning session. They determined who to invite based on which datasets were the most popular, contacted potential attendees, and hosted an orientation session before the event. The CoE also promoted the event to their considerably wider audience and produced large vinyl banners to decorate the space. The City’s then-chief information officer, Chris Moore, a well-known open data advocate in his own right, agreed to be our introductory speaker.

<table>
<thead>
<tr>
<th>Time</th>
<th>Event Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00</td>
<td>Hackathon registration at EPL Makerspace</td>
</tr>
<tr>
<td>9:30</td>
<td>Opening remarks</td>
</tr>
<tr>
<td>9:45</td>
<td>Speed data-ning: participants get to meet various city business representatives</td>
</tr>
<tr>
<td>10:30</td>
<td>Project idea sharing: participants outline what projects they will be working on for the rest of the day</td>
</tr>
<tr>
<td>11:00</td>
<td>Hacking, making, analyzing, writing and all other forms of open data work</td>
</tr>
<tr>
<td>12:00</td>
<td>Lunch is available; participants are welcome to take a break or eat while they continue their work</td>
</tr>
<tr>
<td>4:30</td>
<td>Debrief: Participants share what they have accomplished and learned during that day, share next steps and commitments</td>
</tr>
</tbody>
</table>

Figure 1. Program for International Open Data Day Hackathon, February 22, 2014.

**Tech Infrastructure**

One of the key concerns we identified was that the strength and speed of EPL’s Wi-Fi was of critical importance to the day’s success. We anticipated that the event would impose up to 50 participants actively using the Wi-Fi on top of the regular library usage. To address this concern, we provided the Library’s network specialist with a month’s advance notice of the event. The advance notice allowed him to set up and test a dedicated wireless network for the event. We were also able to take advantage of the resources in our Makerspace to improve Internet access. The Makerspace hosts four desktop PCs and four desktop Macs. We used a switch to connect the desktops to the Makerspace network, giving participants an alternative to using the Wi-Fi. Additionally, we asked registered participants to arrive with any software they planned on using pre-downloaded to their laptops. On the day of the hackathon, we shared the hackathon-
specific network password with participants, and there were no troubles reported with the Wi-Fi.

Cost

McArthur, Lainchbury, and Horn identify venue and food as the two main expenses associated with running a hackathon. The Makerspace proved to be an ideal space for hosting a hackathon because it was designed to be flexible. All the tables in the room were on wheels, allowing participants to organize the space according to their needs. Retractable power cords have been hung from the ceiling to facilitate access to power from most areas of the room. With this infrastructure in place, the only expense was providing coffee in the morning and pizza at lunch. Door prizes were gift certificates for free print jobs on the Makerspace’s 3D printers and Espresso Book Machine, and the “best idea” prizes were library-branded water bottles and coffee mugs.

City Business Representatives

The city invited the staff members responsible for publishing the 10 most popular datasets. None of the city staff had heard of speed data-ing, and they had a number of questions about what they were being asked to do. To answer these questions and make staff feel comfortable with the project, the CoE IT department hosted a catered lunchtime orientation session two weeks before the hackathon. The session was attended by six business representatives and a librarian. By covering the basics—what is a hackathon, what is speed data-ing, why put on this event—and answering specific questions, we were able to relieve confusion and generate enthusiasm amongst the city staff.

Media

A publicity campaign for the hackathon was an opportunity not only to promote attendance, but also to promote the open data movement within Edmonton and to reach out to the international open data movement and join voices with theirs. The CoE IT department took the opportunity to promote the event amongst their own staff and increase understanding and support for its open data initiatives. The CoE and EPL coordinated our promotional strategies to begin in the first week of February. The CoE sent out press releases and tweets on the same day that EPL shared the event logo and a link to the online registration form on our homepage’s carousel. The response was immediate, with 12 registrants in the first 24 hours. In the month leading up to the event, promotion continued with the CoE placing an article in an internal newsletter and mentioning the event on the radio, and EPL sharing tweets and Facebook posts that linked back to the registration page.

The organizers of International Open Data Day also offered opportunities to promote the event in connection with international activities. Open Data Day is primarily coordinated through the use of a wiki site set up for each year. EPL was able to set up a wiki page describing Edmonton’s event (“Edmonton2014”) and add it to a map of events happening around the world. There was debate on the Open Data Day listserv about
whether to use the hashtag #odd2014 or #opendataday combined with the city name. EPL went with #odd2014, but following either provided an exciting view of international events.

**Registration**

Due to the constraints of the physical space, we capped the total number of participants at 50 people. We gave interested participants the opportunity to reserve their spots using an online registration form, which asked for the participant’s name and email address. It also gave them an opportunity to list datasets they would like the CoE to publish. This allowed us to track the popularity of the event and the interests of participants, and make an accurate judgment about catering numbers. The registration form page was the central source of information for interested people and was linked in all promotional activities.

We put the dataset requests anonymously on the event wiki page and encouraged participants to review the page in a reminder email a week before the hackathon. We also let participants know which city business departments would be represented and encouraged them to think about what they could ask. Two registered participants took advantage of the wiki page to share their project ideas in advance of the hackathon.

Two days before the hackathon, we had reached our maximum number of registrants and had seven people on our waiting list. However, the day before the hackathon, a number of people emailed to cancel their registration, and most people on the waiting list had made other plans. Similar circumstances were reported by Eaves in his review of the City of Vancouver hackathon, and he encouraged new hackathons to anticipate “no shows” in their planning. On the day of, we had 44 registered participants, 29 of whom attended the event.

**Open Data Day Hackathon**

The absence of a few participants did not end up deterring from the energy of the event. Enthusiasm had spread amongst CoE employees, and nine attended to take part in speed data-ting; a local open data enthusiast also volunteered to photograph the event. We had reporters from two major local media outlets recording our introductory remarks and interviewing participants.

The speed data-ting session began with the MC introducing participants to the city business representatives and explaining the premise. There were to be three rounds of ten minutes each. The room was buzzing with conversation immediately. At each ten-minute mark, the MC announced it was time to move on, but the microphone was barely heard over the din of the crowd as the conversations continued. Though the activity did not run as anticipated, it achieved its purposes of building connections between citizens and city staff members and promoting discussions about open data.

Following speed data-ting, participants had an opportunity to share the projects they were interested in working on. From there, everyone broke into groups and began
plotting out how they would approach the challenges they had set for themselves. This was the time for organizers to step back and allow participants to make the day their own. Lunch arrived at 12:30 p.m., and in most cases participants brought their pizza back to their laptops and continued to code as they ate. As planned, the afternoon consisted of more coding and project development. Around 4:30 p.m., participants were called to assemble and share what they had achieved.

After half a day of coding, it is not reasonable to expect participants to have created full software programs, but sharing ideas, progress, and even challenges is an opportunity for everyone to learn, build connections, and provide support. Many city staff stayed to watch the presentations. Two participants presented visualizations of city data; the first tracked the energy production of city-run solar panels, and the second playfully mapped which dog breeds were the most popular in Edmonton based on pet registration data. A group of seven participants had created a mock-up of an open data catalogue interface designed specifically for teachers interested in integrating open data into their lesson plans. The winner of our “best idea” prize, as voted by the participants, was a project called Tree Time, which maps all of the city-maintained trees and provides species information.

**Evaluation**

Following the hackathon, participants were sent a twelve-question Survey Monkey questionnaire. The survey was open for two weeks, and, of the 29 participants, 15 responded (51% response rate). This survey was an opportunity to evaluate the event, learn more about the open data community, and collect suggestions for how EPL could improve the support for the community. Additionally, the CoE asked us to gather recommendations for new datasets to publish. Five months after the hackathon, participants were sent a follow-up survey that sought information on the progress of projects started at the hackathon. Participants who had continued to develop their project were asked four questions about their work; participants who had not continued, and who still hoped to continue, their projects were asked what had prevented them from continuing. The follow-up survey was open for two weeks and of the 29 participants, 9 responded (31% response rate).

**Participants**

In the first survey, we asked four multiple choice questions to capture a portrait of the day’s participants. The majority of respondents were between the ages of 18 and 45 (67%); there was a significant number of respondents over the age of 45 (27%), leaving only 7% younger than 18. This could be because the event was not specifically marketed to teenagers. The survey did not ask about gender, but based on registration data we know that a total of 7 women registered for the hackathon, and 2 participated. The majority of respondents had never attended a hackathon before (67%), and respondents’ computer programming skills spanned the spectrum from beginner (40%) to intermediate (33%) to advanced (27%). For the majority of respondents (67%), this was the first event they had ever attended at EPL.
Open Data Day

The three questions asking participants to evaluate the event itself were a mix of open-ended and multiple-choice. Responses to the open-ended question “Overall, what did you like most about the Hackathon?” were very consistent: 73% of respondents mentioned that they enjoyed the opportunity to network with city representatives and people with similar interests. Respondents also mentioned that they liked the opportunity to learn more about open data, and they liked seeing the final products at the end of the day. The suggestions for how we could improve future hackathons were more diverse. A common thread was the suggestion to provide more opportunities for people with similar ideas and differing skill sets to connect with each other. Some specific recommendations were to provide a more detailed agenda and more communication tools like whiteboards (33%). Respondents suggested that we try to tie the hackathon to other events happening internationally by providing examples of completed projects and internationally available datasets (20%). Other respondents addressed their suggestion to the CoE asking for more granular data, corporate data, or more data generally (20%). Individual suggestions requested more activities for beginner hackers, lessons on coding using Apple products, and having a second round of one-on-one conversations with city data representatives at the end of the day. Overall, the feedback was encouraging and positive, with 93% of respondents saying that if EPL hosted another hackathon in the future, they would try to attend.

Supporting the data community

Looking forward, we asked participants what data they’d like to see from EPL and what other ways the Library could support the open data community. A total of 80% of respondents requested specific library data (See Appendix A). They also suggested some uses of library data in library services, such as offering indoor mapping tools based on machine-readable shelf locations to help people find items, and making a dashboard of live library data feeds. Consistent with this feedback, 80% of respondents indicated that providing library data is a way that EPL could best support the open data community. Participants were encouraged to select as many strategies for supporting the open data community that they felt would be meaningful from a multiple-choice list. Other popular strategies included: hosting more hackathons (67%); promoting apps, visualizations, and games made from open data on epl.ca (67%); and offering code camps and other courses on coding (67%). Less popular, but not-insignificant, strategies were sharing information about open data over social media (47%), and promoting books, e-books, and e-audiobooks about open data in our collection (47%).

Hackathon products

The first survey asked participants whether they were going to continue to develop the projects they started at the hackathon. The majority of respondents said they definitely would continue to develop their projects (40%); 27% said they probably would, and 33% were unsure. No one selected the “probably will not” or “definitely will not” options. This could reflect a correlation between a participant’s likeliness to fill out a post-hackathon survey and enthusiasm for their project. The five-month follow-up survey paints a
different picture. Of our nine respondents, two had continued to work on their projects (22%); four had not continued (44%), and three had not continued but hoped to continue in the future (33%). Of the respondents who had not continued to develop their projects, 75% gave “I lost interest” as a reason, and 25% gave “I encountered a problem I could not fix” as a reason. Most of the respondents who still plan on continuing to develop their projects gave “I didn’t have enough time” as a reason (67%), and 33% listed “lack of data” as a reason. How to provide post-hackathon support for app development is one of the challenges the Library and the CoE will tackle for International Open Data Day 2015.

Promoting open data

Both the first and second follow-up surveys asked for examples of positive stories about the hackathon or hackathon products that could be shared with the general public. The benefit of collecting stories like this is that they can be used to promote open data and future hackathons. As feedback from the first survey suggested, examples of completed projects can also help give new civic hackers some direction. We received seven stories about the hackathon in the first survey. Most reiterated what a positive experience it was to meet people with similar interests, and to meet city staff who provide the data they use. One story in particular tied together the exciting potential of open data and the benefits of hosting a hackathon:

I am working on a project that could potentially save lives through predictive analytics, and through discussions with people from various professional backgrounds, I found a solution to an issue I was having with the data. It took bringing all the different people and their experience together to make it work. Thanks.

In the second survey we asked for permission to share completed works or works in progress. The creator of Tree Time was happy to share the work he had done as an example of a civic hacking project.¹

The Open Data Day hackathon played a significant role in supporting the open data community in Edmonton. A September post on one of Edmonton’s most prominent blogs declared that “Open Data in Edmonton is exciting again,” crediting the Library’s hackathon for having “kicked things off” (Male). A second hackathon titled HackYEG, organized by a community group, was planned to coincide with the American National Day of Civic Hacking and was also hosted at the Library. Attendees of both hackathons have since formed a citizen-led open data advocacy group called Open Edmonton, which held its first meeting at the Library and personally invited the digital public spaces librarian to take part. Open Edmonton meets regularly and has become an active voice for the open data movement in Edmonton.

¹ At the time of publication, Tree Time was accessible at http://treetime.ca/edmonton/
**Discussion**

Hosting the hackathon gave EPL a sense of the demographics of the open data community and their needs. That the majority of participants had never attended a hackathon before, and listed their programming skills as beginner, suggests that participants held a general interest in open data and trusted that the Library could function as an introductory support venue. While the type of project development that we encouraged at the hackathon could benefit from the contributions of people with many different skills and skill levels, the openness of the schedule appeared to intimidate some participants. To make civic hacking events more inclusive to beginners, a number of organizations have devised simple projects that share the value of openly accessible information and collaboration. For example, organizers of the National Day of Civic Hacking proposed the Open Data Challenge and the Digital Front Door/Digital Divide projects, which encourage citizens to evaluate their city’s open data catalogues and online information portals, respectively, and offer suggestions to improve accessibility and quality (“Open Data,” “Digital Front”). For the same event, an OpenStreetMap edit-a-thon was held at Cleveland Public Library to crowdsource the creation of openly available local data (“OpenStreetMap”). Future hackathons, including next year’s EPL event for International Open Data Day, could incorporate beginner activities, or the Library could run them throughout the year.

As an effort to support data literacy, the hackathon was primarily an opportunity for participants with existing data literacy skills, such as how to find, interpret, and manipulate data, to practice them in a social environment. The speed data-ing activity was a unique opportunity to learn about the production process of the City’s public datasets. After discovering the data literacy core competencies defined by Calzada-Prado and Marzal, it was determined that the speed data-ing activity supported competency 1.2: “be aware of the role of data in society, how they are generated and by whom, and their applications, as well as the implications of their use” (130). The request for the event to be more inclusive of beginner programmers points to the need for a more thorough understanding of data literacy competencies. While they were not used in developing the hackathon programming at EPL, Calzada-Prado and Marzal’s breakdown of data literacy core competencies can help staff develop and assess meaningful data literacy programs in the future.

Hosting the hackathon also put EPL in a strong position to begin releasing meaningful internal data. The Open Knowledge Foundation’s *Open Data Handbook* establishes five criteria for determining which data to prioritize sharing: privacy, community priorities, cost, ease of release, and peer choices. With the follow-up survey results, we have a list of the community’s library data priorities that will direct the development of our open data initiative. Partnering with the City on the hackathon also gave us the contacts to quickly and easily begin releasing data on the Library’s locations, opening hours, contact information, and facilities and Wi-Fi availability, which fall under the final two criteria: ease of release and peer choices.
Conclusion

The International Open Data Day Hackathon at EPL demonstrated the Library’s interest in being an engaged member of the open data community and put us in an excellent position to understand and respond to its interests and needs in the future. The list of recommended datasets we received is directing the development of our open data releases, and the vocal support of the community is an internal motivation. We achieved our goal of reaching out to the open data community in Edmonton, many of whom had never attended an event at the Library before. Particularly gratifying is the local growth of the movement since the hackathon, and the knowledge that the Library is an active member and supporter of a citizen movement that shares the Library’s value of open access to information.

Works Cited


Appendix A – Library Data Requests

The following lists break down the responses to the question “What data would you be interested in seeing from the Edmonton Public Library, if any?” from the Open Data Day Hackathon Survey. The survey was distributed to the 29 hackathon participants who registered at the event and yielded 15 responses between February 24 and March 3, 2014. A total of 12 survey respondents answered this question.

Data requested by Hackathon participants

- Service usage data
  - Program date and time information in an API such as SODA that serves it up as JSON\(^2\)
  - Volumes of people in the branches
  - Types of usage
  - Daily data
    - Branch visits
    - Computer logins
    - Website logins
    - Catalogue searches
    - Interlibrary loan requests
    - Hold requests
    - Memberships issued by branch
    - Fines levied by branch
  - Catalogue search history
- Collection data
  - Number of books in the collection
  - Types of books in the collection
- Circulation data
  - Most popular books / Book and e-book popularity statistics / trends in book choice\(^3\)
  - “Circulation statistics like Vancouver’s Library [https://www.vpl.ca/opendata/repository/catalogue_2012/]”
  - Circulation by branch / daily books borrowed (each branch) / daily media borrowed (each branch)
  - Checks-in and check-outs by branch / Checkout, renewal and return rates
  - Late fees
  - Delinquency rates

---

2 We already have an events API through our website [www.epl.ca/opendata](http://www.epl.ca/opendata)

3 Similar requests made by different participants have been listed together separated by a “/”