

## **2007 Code4Lib Conference Report**

By: Antonio Barrera, Parmit Chilana, Kevin Clarke, Michael Giarlo

### **Introduction**

[Note: This report was initially published in volume 24, issue 6 of Library Hi-Tech News (Emerald Group Publishing Limited), on pages 4-7. The abstract may be accessed online via the following URL: <http://www.emeraldinsight.com/10.1108/07419050710823247>]

The 2007 Code4Lib conference was held at the University of Georgia, February 27th to March 2nd. The event was the second annual conference for the Code4Lib group, which derives its membership from an e-mail list and an IRC (Internet Relay Chat) channel of the same name. Geared toward computer programmers, the conference revolves around technological problems and solutions in the context of libraries.

This year's conference drew approximately 120 attendees and was sponsored locally by the University of Georgia and the Georgia Public Library Service. Other sponsors included library technology firms OCLC, Talis, LibLime, and Logical Choice Technologies, and the past year's host organization, Oregon State University. The event was organized in a democratic fashion, with community members voting in the open on the location of the conference and the program. The conference began with a one-day pre-conference session on the Solr search system, which was well-attended, and also included two keynote addresses, numerous scheduled talks, and other activities.

Participation among attendees was encouraged via lightning talks and breakout sessions. Lightning talks allowed for over two-dozen additional speakers to participate actively in the conference. Each speaker had a five-minute time limit to describe some interesting technology or give an update on a work-in-progress. Breakout sessions presented an opportunity for individuals to gather in small group settings to discuss a topic of need or interest.

Code4Lib provided an opportunity for library programmers to share their experiences, knowledge, and code with colleagues from across the globe. As Karen G. Schneider, former Associate Director for Technology and Research at the Florida State University Libraries, noted during the opening keynote, libraries must improve their ability to control the content they create or license. One such way is to develop tools to manage and provide access to this content. [5] Code4Lib serves as a means of helping programmers focus on finding solutions to their libraries' needs.

### **Background**

The Code4Lib conference is a relatively new introduction to the world of library conferences but it is grounded in pre-existing communities and traditions. Placing the conference in context requires a look at the broader Code4Lib community and the events and people around which it has coalesced.

The Code4Lib community, as one might deduce, is a group of computer programmers who largely work for libraries. Some come from the IT industry; others are bona fide librarians. All have a common goal of writing, analyzing, managing, tweaking, testing, hacking, or otherwise using code. In 2003, a number of library- and programming-related mailing lists had been created for very specific purposes: one for this programming, another for that markup language, and so forth. A group of library programmers got together in the fall of that year to create a more general mailing list, one that would be relevant to all manner of programming and markup languages, and it would be known as "Code4Lib." [4]

Two catalysts transformed Code4Lib from a small mailing list into an active, international community: a chatroom and Access, the premier library technology conference.

The Access conference, under the aegis of the Canadian Association of Research Libraries, has been at the forefront of library technology since its inception in what was a landmark year for the Internet, 1995, when "the government and the organizations that built [the Internet] ... handed [it] over to commercial networks;" [1] when sites like Amazon, Hotmail, and eBay were created; when the Java language was introduced; when the first version of Internet Explorer was released; and when Netscape went public. Each of these events shaped the future of the Internet and the Web, and the Access conference was well-timed to translate these bold steps forward into how they might solve library users' needs. As early as 2002 [2], the Access conference added a new feature: the Hackfest, where programmers and librarians alike could gather, share ideas about library services, and create them during an intensive (not to mention intensely fun) day of coding, sketching, and other interaction.

Access served as an incubator for much of the early Code4Lib community. Not long after the original Code4Lib mailing list was created, the list community decided to set up a chatroom on the Freenode network of IRC; in a fit of inspiration, they named it #code4lib. The 2004 and 2005 Access conferences provided early Code4Libbers the opportunity to meet face-to-face, and the community grew in leaps and bounds during this period. And the more the community grew, the greater the demand became for more regular meetings than the annual Access offered. [4]

Hard on the heels of Access 2005, the #code4lib regulars decided to try their hand at hosting an Access-like conference in the U.S.; it would need to be informal, single-track, and highly technical. It would also pull a page out of the "unconference" playbook; though attendees would not be required to participate, as is common in "unconference"-style conferences, there would be plenty of opportunities to participate in some capacity or another.[3] Additionally, the Code4Lib conference "was created through a grassroots, open manner, with members of the Code4Lib community working together to plan, organize, and run the conference. Community voting decided everything from the conference location to the accepted presentations and the design of the conference t-shirt." [6] Moreover, much of the discussion and voting was done out in the open, either

on public mailing lists or on publicly available, unrestricted websites, enabling every potential attendee to become an active part of the Code4Lib community even months in advance of the conference. These characteristics, its focus on practical innovation [7], and its commitment to affordability -- the conference registration fee was \$125 -- would set it apart from other library technology conferences.

In February of 2006, just three short months later, the first Code4Lib conference was held in Corvallis, Oregon, where it was hosted by Oregon State University. The format of the conference was similar to that of the Access conferences [8], with sessions divided as follows: fifteen twenty-minute presentations; three rounds of successive five-minute talks, known as "lightning talks," totaling nearly three hours over the course of the conference; two half-hour breakout (or "Birds of a Feather") sessions; and two keynotes, the first of which was, in Code4Lib fashion, delivered via teleconference by the Georgia PINES development team. While the keynotes and twenty-minute presentations were voted on prior to the conference, the lightning talks and breakout sessions were mostly on-the-spot creations, which was highly conducive to the participatory spirit the organizers had intended.

Despite having been conceived and planned over a three-month period, largely by individuals who had little experience in planning and organizing conferences, the initial Code4Lib conference was a hit, drawing over 80 attendees. [7] Some were Access conference regulars; some were members of the growing #code4lib community; and many more were new to the community. The conference drew rave reviews, living up to its billing as "the" event for technologists building digital libraries. [9] Perhaps the biggest question at the end of the 2006 conference was if there would be another in 2007.

One issue that needed addressing after the 2006 conference was the ratio of male to female attendees which though perhaps not disproportionate in technology industry circles was nonetheless skewed for a library conference. While organizers were not necessarily chasing a golden ratio, there was a desire to have a more diverse audience. Oregon State University sponsored one scholarship for a person in a "principal minority group" [11] and another for a female. The recipients of the scholarships, Nicole Engard of Jenkins Law Library (now of Princeton Theological Seminary) and Joshua Gomez of the Getty Research Institute[9], received funds to cover the cost of the conference. Some might criticize this as a small step but it was a step in the right direction, and the community looks forward to similar sponsorships and other methods of ensuring attendee diversity in future conferences.

When voting on the venue for the 2006 conference concluded, Oregon State University came out on top, with the Georgia Public Library Service taking second. It was only natural that they would get to host the second annual Code4Lib conference, this time held on the University of Georgia campus in Athens, GA. Building upon the reputation it had earned a year earlier, the 2007 conference attracted approximately 75% more attendees [10] without losing the informal community feeling; the pool of proposed talks grew to fifty [4], of which sixteen were chosen; and a pre-conference workshop was added.

## **Major Themes**

This year's conference covered a range of topics with many interesting presentations, lightning talks, and break out sessions. One of the purposes of the Code4Lib mailing list and IRC channel is to provide a shared participatory environment where people can learn from others' works in progress. The conference did not disappoint those who have come to expect this type of experience.

Like one would expect, the presentations at the conference were geared towards a technical audience. They consisted mostly of demonstrations of practical applications with introductions to relevant emerging technologies. Challenges that were addressed throughout the conference included:

1. Meeting the expectations of users familiar with Web 2.0 sites like Google, Netflix, Flickr, and Amazon,
2. Securing institutional resources for the development of library applications, and
3. Learning new tools and programming languages in a rapidly changing environment.

Of the many themes explored at Code4Lib 2007, there seemed to be three major ones that resonated throughout the presentations, lightning talks, and break out sessions. The first theme was new developments related to online public access catalogs (OPACs) and integrated library systems (ILSes). Schneider, the opening keynote speaker, spoke about "Five things we can fix" leading her to the observation that the open source software Evergreen, with some minor marketing problems, is very well positioned in the ILS market. [12]

## **OPACs and ILS**

The topic of OPACs and ILSes stirred discussion at this year's conference. Desires for improvements in the existing features of our current ILSes were mixed with issues of use and usability, our patrons' demands for newer social software like features, and the high price tag of many of the more innovative retrieval systems. Approaches discussed by presenters and conference attendees ranged from layering new services over existing systems to replacing existing systems, or pieces of existing systems, with newer open source alternatives.

For instance, several of the presentations at the conference discussed working new features like faceted browsing, social bookmarking, and improved search results through subject recommendations into our existing OPACs. Others experimented with the replacement of OPACs by XML-based technologies and standards like MARCXML. Adding web services over an existing catalog was also an idea explored by several of the presenters. ISBN REST requests, an RSS layer for all catalog searching and browsing activities, and OpenSearch integration with the library website's search were discussed as options.

## Searching

The second theme was searching. Erik Hatcher, in his day two keynote session, took up the issue of searching and, in particular, faceted searching in an attempt to demonstrate how libraries and librarians could be taking advantage of more sophisticated search tools. Hatcher, who literally wrote the book on Lucene [13], led the pre-conference, an intensive workshop on the Solr opensource search engine. The pre-conference itself attracted as many attendees as the 2006 conference [9], and was a timely addition to the conference, as a number of Solr- or Lucene-based library search engines have been cropping up over the past couple years. And the best has been saved for last: the pre-conference was free.

There is clearly a growing demand for more effective searching (and finding) of library resources. This was reflected in the significant role in the presentations and lightning talks given at the conference. What was interesting, perhaps, is that most of the search tools demonstrated or discussed at the conference were open source. On the other hand, this isn't that surprising given the technical and do-it-yourself nature of the Code4Lib group.

Search applications discussed at the conference ranged from hybrid federated search services coupled with knowledge-base management, caching, and OpenURL integration to faceted, XML-based web services built on top of successful search engine libraries like Lucene. Some presentations were pragmatic; for instance, one explained how search results could be improved by integrating smart subjects and log reviewing into the process of creating a search application. Others were more forward thinking, like one that discussed merging metasearching and resolver user interfaces with technologies like OpenSearch and Zeroconf.

## New Technology

One of the main purposes of the Code4Lib conference is to expose attendees to new standards and technologies. The third theme of the conference was the ongoing experimentation with new technologies and techniques that many Code4Libbers participate in on a daily basis. At this year's conference, standards like the Atom Publishing Protocol and XQuery were explained and put in the context of library application development. New technologies, like Ruby on Rails, were also discussed. One Ruby on Rails application, BibApp, was well received as a mashup of article databases, citation management software, and digital repositories.

The importance of application programming interfaces (APIs) was also discussed. The need for library specific APIs that allows developers to access programmatically our data stores is a common theme in the Code4Lib mailing list and IRC channel (and elsewhere in the library world). Many of the presenters also spoke about reducing the barriers to developing and using better library systems. Two successes include the Library In a Box project and an example from a public library system that has deployed Groovix as a thin-client solution.

## **Beyond the Themes**

While the Code4Lib conference features a set group of presentations whose topics are selected by the Code4Lib community, there is also room for many more spontaneous presentations. These shorter lightning talks and breakout discussions were given the same status in the program as their longer, scheduled siblings. The diversity of topics represented by the lightning talks and breakout sessions is really too great to try and briefly pin down, but what is noteworthy is that many of them have taken on a life of their own.

One of the lightning talks discussed how XForms is being used in one digital library project, and spawned during the course of the conference, a wiki on the Code4Lib site where the discussion has been continued even after the conference's end. Similarly, experiments with Solr, the web services search application featured at the pre-conference workshop, have also continued in several of the libraries, which sent people to the Code4Lib conference. Beyond the discussion of topics or networking with one's librarian and programming peers is the practical sharing and building of projects that continues even after the conference has ended.

Those who were unable to attend the 2007 conference might be interested to know that audio and video are available for all of the talks, for free, over the web.

## **Conclusion**

The 2007 Code4Lib conference provided knowledge and encouragement to the programmers, developers and coders from libraries. Many in the Code4Lib community learn from each other either via e-mail, blogs, and chat. This conference gave these programmers an opportunity to meet face-to-face and continue learning from one another, networking, and building community.

Planning for the 2008 Code4Lib conference has begun in earnest. Three institutions submitted proposals to host the third annual conference and Portland, Oregon edged out Madison, Wisconsin and Vancouver, British Columbia. Exact conference dates and the location are yet to be announced, but the planners do expect to have these details pinned down by mid-summer, 2007. Readers interested in tracking Code4Lib 2008 conference developments may wish to watch the website that's been set up for the conferences. Any readers that may be interested in getting involved are invited to join the planning committee. All that is required is joining the code4libcon Google group and speaking up; the rest is up to the democratic process.

Beyond the 2008 conference, the sky's the limit for the Code4Lib community. It may continue running annual conferences, or 2008 may be the last; it's entirely up to the community. Interest has also been shown in a Code4Lib Europe conference. Though no concrete decisions appear to have been made, organizers have expressed the desire to make sure the European conference does not conflict with either the Access conference or the Code4Lib conference. Discussions are ongoing to create a peer-reviewed Code4Lib

e-journal, and anthologies are being compiled of the best library technology blog postings from 2006 and 2007.

As with any conference, the event is greater than the sum of its planned sessions. The networking presented an opportunity for individuals to meet birds of a feather. As evidenced by the themes and discussions at the Code4Lib conference, library programmers are willing and able to fill the gaps in software to help their libraries innovate and compete for 21st-century mindshare. Open-source and local development can provide unique value to the service libraries strive to provide. There will be success and failures along the way, but conferences and communities such as Code4Lib provide programmers a sense of empowerment to continue to improve on the status quo.

## References

1. Ruthfield, S 1995, *The Internet's History and Development*, viewed 28 June 2007, <http://www.acm.org/crossroads/xrds2-1/inet-history.html>
2. Chudnov, D 2005, *Access Hackfest History and Links*, viewed 28 June 2007, <http://old.onebiglibrary.net/yale/curtis/hackfest/>
3. Ohio Public Library Information Network 2007, *OPLIN 4cast Podcast #16: Code4Lib Conference Report*, viewed 28 June 2007, <http://www.oplin.org/4cast/index.php/?p=67>
4. Chudnov, D 2007, 'Code4libcon shows what a participatory conference looks like', *Computers in Libraries*, vol. 27, no. 3.
5. Schneider, K G 2007, *Code4Lib Keynote, Again*, viewed 28 June 2007, <http://freerangelibrarian.com/2007/03/03/code4lib-keynote-again/>
6. Chudnov, D and Frumkin, J 2006, 'Code4lib 2006', *Adriane*, vol. 47 viewed 29 June 2007 - <http://www.ariadne.ac.uk/issue47/code4lib-2006-rpt/>
7. Schneider, Karen 2006, *Code4Lib Conference Notes Up*, viewed 29 June 2007 - <http://freerangelibrarian.com/2006/02/18/code4lib-conference-notes-up/>
8. Datema, Jay 2007, *code4lib 2007*, viewed 29 June 2007 - <http://www.bookism.org/open/2007/03/13/code4lib-2007/>
9. Schneider, Karen 2005, *Code4Lib 2006!* viewed 29 June 2007 <http://freerangelibrarian.com/2005/12/02/code4lib-2006/>
10. Chudnov, D 2007, *code4lib 2007 brain dump*, viewed 29 June 2007 < <http://onebiglibrary.net/story/code4lib-2007-brain-dump> >
11. Frumkin, J 2007, *Scholarships for the 2007 Code4lib Conference*, viewed 29 June 2007 - <http://digitallibrarian.org/2007/1/2/scholarships-for-the-2007-code4lib-conference>
12. Schneider, Karen 2007, *Hurry Up Please Its Time*, viewed 29 June 2007 <http://freerangelibrarian.com/presentations/hurryuppleaseitstime.pdf>
13. *Lucene in Action*, Manning Publications, US, viewed 29 June 2007 <http://www.manning.com/hatcher2/>

## Links

- Solr Search Server - <http://lucene.apache.org/solr/>

- Evergreen ILS – <http://open-ils.org>
- Access Conference - <http://access2007.uvic.ca/>
- Code4Lib Conference - <http://code4lib.org/conference>
- 2006 Code4Lib Conference - <http://code4lib.org/2006>
- 2007 Code4Lib Conference Schedule - <http://code4lib.org/2007/schedule>
- Code4Lib Conference Planning Group - <http://groups.google.com/group/code4libcon/topics>
- XForms Wiki - <http://xforms.code4lib.org/>
- Atom Publishing Protocol - <http://www.ietf.org/html.charters/atompub-charter.html>
- OpenSearch - <http://opensearch.org>
- BibApp - <http://code.google.com/p/bibapp/>
- Zeroconf - <http://www.zeroconf.org/>
- XQuery - <http://www.w3.org/XML/Query/>
- Ruby on Rails - <http://www.rubyonrails.org/>
- REST (Representational State Transfer) [http://en.wikipedia.org/wiki/Representational\\_State\\_Transfer](http://en.wikipedia.org/wiki/Representational_State_Transfer)
- Groovix - <http://groovix.com/>
- Pines - <http://gapines.org/>
- OpenURL - [http://www.niso.org/standards/standard\\_detail.cfm?std\\_id=783](http://www.niso.org/standards/standard_detail.cfm?std_id=783)

Antonio Barrera ([abarrera@princeton.edu](mailto:abarrera@princeton.edu)) is the Library Web Development Manager at Princeton University Libraries, Princeton, NJ, USA; Parmit Chilana ([pchilana@princeton.edu](mailto:pchilana@princeton.edu)) is a Digital Library Specialist in the Systems Office at Princeton University Libraries; where Kevin Clarke ([ksclarke@princeton.edu](mailto:ksclarke@princeton.edu)) is Digital Projects Programmer; and Michael Giarlo ([mgiarlo@princeton.edu](mailto:mgiarlo@princeton.edu)) is also a Digital Library Specialist.