Hackathons at Berkeley (HackFSM and HackTheHearst)

Our next Research IT Reading Group will focus on campus hackathons. Last year, Research IT supported two hackathons: #HackFSM for the Bancroft Library’s Free Speech Movement Digital Archive and #HackTheHearst for the Phoebe A. Hearst Museum of Anthropology. Mary Elings, Head of Digital Collections at the Bancroft, and Michael Black, Head of Research & Information Systems at the Hearst Museum, will present on their experiences as hackathon organizers. We hope to continue tinkering with the format of the event and explore hackathons as a place of mutual benefit for a variety of students, community members, and hosting campus organizations. We heartily invite the perspectives of other campus members who have hosted, participated in or supported hackathons.

Some questions up for discussion will be:

- How can we make hackathons a beneficial experience for both participants and hosts?
- How do hackathons in the GLAM (galleries, libraries, art galleries, museums) space compare with hackathons in the private sector / tech industry?
- What constitutes a compelling challenge for hackers? What is reasonable scope for a hackathon project?
- What preparation goes into a hackathon? (e.g. data prep, API documentation, event planning, outreach, co-sponsors, etc.)
- What resources can / should be made available to participants? What criteria should be used for judging?

When: Thursday, January 29 from noon - 1pm
Where: 200C Warren Hall, 2195 Hearst St (see building access instructions on parent page).
Event format: The reading group is a brown bag lunch (bring your own) with a short ~20 min talk followed by ~40 min group discussion.

Quinn Dombrowski (Research IT) and Camille Villa (Research IT) will be facilitating. Mary Elings (Digital Collections, Bancroft Library) and Michael Hearst (Hearst Museum) will be presenting on their experiences organizing HackFSM and HackTheHearst, respectively.

Please read/review the following in advance of the 1/29 meeting

=====================================

==> “Hackers Descend On a Campus Near You”, Chronicle of Higher Education

==> “#HackFSM: Bootstrapping a Library Hackathon in Eight Short Weeks” (skim, but focus on sections 4.1 and 5)

=====================================

Attending:
Scott Peterson (Library)
Chris Hoffman (Research IT)
Mary Elings (Bancroft Library, Digital Collections)
Camille Villa (Research IT)
Sherry Lochhaas (Hearst Museum)
John Lowe (Research IT)
Michael Black (Hearst Museum)
David Fulmer
Patrick Schmitz (Research IT)
Sharon Goetz (Bancroft Library - Mark Twain Papers)
Steve Masover (Research IT)
Aaron Culich (Research IT)
Jessica Smith (Angel Hack)
Troy Petronof (Angel Hack)
Rick Jaffe (Research IT)
David Greenbaum (Research IT)
Aron Roberts (Research IT)
Bill Allison (IST)
Quinn Dombrowski (Research IT)

Presenters:
Mary Elings
Michael Black
John Lowe

Notes:
Mary Elings on HackFSM

HackFSM: An interdisciplinary DH hackathon, April 1 - 12, 2014
Free Speech Movement Digital Archive

13 teams of UC Berkeley students tasked with designing a new interface
FSMDA interface built into the 1990s, one of the Bancroft’s first digital projects
8 weeks from idea to launch
Library’s first hackathon
impetus - 50th anniversary of the FSM in the fall, something to highlight the collection and give it a fresh look
First challenge: explaining to Library staff what a hackathon is
we tried: a positive, creative event putting something interesting together in a short period of time
success story: Janine Heiser, an I School student now working on a collections tool at the Bancroft as a Fellow now, who participated in the hackathon

Goals:

- humanities focus
- team composition requirement to emphasize interdisciplinary focus: at least one “hacker” and one “humanist” on each team. not just engaging with the content but engaging with one another
- a longer coding period: 12 days so that people could work out communication issues. (Though at 12 days...I guess we’re supposed to call it a web app challenge)
  - 24 - 72 hour hackathons not as appealing to women, tried to offer healthier food
- DH at Berkeley had been focused on faculty and grad students; this was an opportunity to reach out to undergrads
- first time exposing a collection via API; one of our goals in the near future, the hackathon was a wonderful chance to try this out

What we provided:

- Apache Solr API managed via IST’s API Central
- Piazza as a communications platform — with each other, with mentors, with organizers
- kickoff event at the library to mingle, eat, start coding
- closing event on Cal Day at the I School open to the public, present final projects
- during the event: workspace available in the D-Lab
- Prizes - 1st place: macbook air, 2nd place: iPad minis

How did it go?

- 13 teams, 8 presented
- winner was pretty clear
- great crowd for UL, university archivist, associate dean for social science was the judge
- FSMers actually came to the event, donated content there, and will hopefully include more in the future
- DynaWeb...static funky site  a fully deployable site

Post hack:

- good:
  - student feedback:
    - took part b/c the content was interesting, not just the prizes.
    - learned more about the FSM
  - publicity:
    - were on teh main page of the library
  - goals met:
    - 8 weeks
    - some students provided simultaneous editing environments, research dashboards...unofficially, we had 25% women participants
    - successful imitation by HTH
- bad:
  - didn’t now there’s a $600 limits on prizes, had to be rerouted as a scholarship. lots of red tape
- ugly:
  - maintaining winning sites
  - were lucky the 1st place site was a full deployable site
  - 2nd place site only got deployed (shakily) recently
student engagement! our researchers and faculty of the future, more and more of this collaboration will be in the works in the future also great to see inter-institution collaboration

**Michael Black on HackTheHearst (HTH):**

I had wanted to have a hackathon at PAHMA for a couple years. We had seen great things going on with culture hackathons in the European cultural sector (e.g. Open Culture hackathons, Rijksmuseum in the Netherlands) marketed as development of open source museum collection tools aside from HackFSM perspectives, also followed the advice of AC County 2012 hackathon

Lessons learned + comparisons with HackFSM

- need more than 2 months to pull of a hackathon...went for 4 months.
- most of our time was spent trying to find donors and sponsors for the site
  - stroke of luck: an EMC2 employee had an undergrad degree in Anthro and saw our call for help in an alumni newsletter
- tried to work on prizes that didn’t have financial value but most were experiential: meet with archivists, take tours
- participants didn’t seem to care about prizes; participants were actually surprised that there were prizes
- press, mentors, and judges were more concerned about prizes
- HackFSM had trouble finding digital humanists; we ended up with the opposite problem: we had a no-show for EECS students

FSM: participants didn’t feel like they had enough clarity on desired outcomes
HTH: tried to define several competition tracks; opening program with target users; drop-in sessions (no-shows for about 50%); ...had too much info on websites
removal size and team composition requirement
getting lots of n = 1 teams, and need to encourage inclusivity
winning team ended up being 11 people, which seemed a bit unfair

FSM: mentors were underutilized
HTH: tried to emphasize how important mentors will be to team/project success. personal introductions, etc. but still mentors were barely taken advantage of

The projects:

- 9 prize categories, but not all prizes were awarded (i.e. there weren’t enough competitors in every category)
- Yapi Kapi: map-based app allowed K-12 users to select a portion of their state and find artifacts and cultures from that area
- Time Capsule: Ancient Egypt: rich focus on one artifact; team is working on developing more lesson plans
- Museum Manager: 2nd place in K - 12 S objects exhibit building game, dating objects

Outcomes:
a success but less well-attended than we had hoped. we had expected 250 - 300 people. had 200 pre-registered plus drop-ins, but we ended up with only 103
used Eventbrite for registration; had a few ticket options for people who could afford to donate a little money, raised about $200
we spent 4 months on fundraising and up to the last minute, we were worried we wouldn’t have funds to throw the event. we were $200 in the black at the end. we approached 220 orgs, cold calls and letters — but mostly it was personal connections that resulted in donors. In the future: would recommend having 2+ volunteers dedicated to finding donors and sponsors

Venues were difficult to find/ expensive: few options for >150 attendees (and suitable for hacking)
finding judges was difficult — perhaps the API creator should not have been a judge (was less accessible)

publicizing the event: website, facebook page, PAHMA twitter, posters, student ambassadors to classrooms, presented at Bay Area DH, UCB News Center, Daily Cal, front page of the UCB website 2 weeks prior...yet still didn’t have full awareness w/in institutions

having no idea how many participants we’d have on the day of the event made it extremely difficult to plan: venue, rentals, planning food (60% of budget), publicity, finding sponsors (#1 question from potential sponsors: how large is your hackathon? anything under 500 is considered a small hackathon)
almost no undergraduates showed up: bad choice to host a hackathon at the beginning of the semester? competition with other big ticket hackathons happening around then (but not simultaneously)?

held back API keys until people were in teams...but they should have been distributed more widely

too many rules and requirements:

apps that should have been disqualified that but weren't

prizes that should have been awarded but weren't

plagiarism overlooked and even rewarded

judges didn't fully understand award categories

campus native community in uproar over publicity of winning app

why weren't we invited? why weren't we consulted? why weren't there native speakers? why weren't there native mentors? we'll need to find a better way
to show this is relevant to humanities folks

had to forego post-event publicity because of Yapi Kapi issues

ideally we'd like to vet apps before they're available to / announced to the public to make sure we avoid things like this

HTH 2015 - should there be one? when? how to increase undergrads?

how do we fund it? can't dedicate 4 months to finding donors...partner with campus departments and orgs? perhaps align with a class?

John Lowe on lessons learned from surveying judges and mentors

41 members on the list, 3 responses, chased down some folks for interviews
domain judges and technical judges

most of the technical judges had experience with hackathons, but noted that academic setting 2 wk hackathon is a very different usually hackathons are a set of experts in some technology showing up to an event to build an idea
objective needs to be clear: can't expect a finished app, deployable, meeting security reqs, etc...need to be realistic in the future prizes may or may not be important
heavy event attrition — many will show up but few will follow through
steep learning curve to get familiar with an API, codebase...how to call an API, structure of the data, content of the API, becoming familiar with collection information in the anthropological space
messaging is important: release of results, managing communication with both involved and non-involved parties, consensus and transparency, judges were confused about what they were and weren't supposed to be doing

---

Discussion

Jessica, AngelHack:

100 person hackathon is a huge success! esp. considering how niche and specialized your hackathons they are. we would usually go for around 80 people. 500 person (i.e. CallHacks) hackathons are huge!

Michael: we actually only had 8 teams at the end

Jessica: [background on AngelHack] we usually have 100 people, 20 teams
AngelHack global hackathon org, started out hosting 1, 11 35 50 hackathons this year! host weekend long hackathons and also do virtual hackathons, we have an accelerator program: winning teams put through accelerator, mentorship for 12 years, flown out to Silicon Valley to present, pitch, and demo to 100s of investors. had 2 acquired by Google, one company has raised over 10 million (Osper, an app for giving your children an allowance, monitoring it)

Aaron: Data Carpentry workshops emerging out of Software Carpentry

Mary: the API was a pain for the library. we have all this great content but...

we're thinking, small curated groupings of things. Quinn's helping me with a project on our Tebtunis Papyri collection...may work with Michael, b/c the Hearst also has Tebtunis papyri materials

Look forward to a Wikipedia edit-a-thon on women in history , esp. early California women who ran most of the ranches, husbands heading out for gold etc, the Phoebe Hearst papers

Camille: what if we broke up these hackathons into two events: a "data-thon" to practice data wrangling, analysis, viz and get familiar with collections data, and then a hackathon for building web apps? D-Lab hosted a Datathon with the American Sociological Association, should ask them how that went

Additional comments

From Lara Wolfe, in absentia:

* Visibility has proved to be the biggest driver for our hackathons. We receive greatest turnout when we offer the opportunity to have worked displayed in public places or used in new situations. This does not only apply to students, but to artists, coders etc., who would like to increase their visibility. We'll also offer an additional small Amazon gift card prize of $100 or so to each of the winning team members, but predominantly we focus on giving their work a platform.

* We've had great success with turn-out by including a "star" mentor/judge who is there throughout the day. For example, Casey Reas, who created processing.org brought in 75 people to SF for a one day hackathon. This year, we're bringing in artists Enrique Mayorga and Carlos Martinat. Participants want the opportunity to work closely with these individuals.

* We're learning that our hackathons are most successful when we are clear in our aims and the outcomes we desire. We've had some duds when the mandate was too vague and the projects produced were ultimately all over the show.