The First Workshop on Hacking and Making at Time-Bounded Events

Anna Filippova, Erik Trainer (Editors)

CMU-ISR-17-104
August, 2017

Institute for Software Research
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213

This work was supported by Grant Number G-2017-9680 from the Alfred P. Sloan Foundation, James Herbsleb, PI.

Keywords: time-bounded collaborative events; workshop; hackathons; sprints; codefests; radical collocation; community building; innovation; collaborative learning, cscw.

Abstract

Time-bounded collaborative events in which teams work together under intense time pressure are becoming increasingly popular. In 2015, collegiate hackathons alone attracted over 54,000 participants across 150 events. While "hackathons", that is, competitive overnight coding events, are one of the more prevalent examples of this phenomenon, there are many more distinct event design variations for different audiences and with divergent aims: "sprints" bring together existing communities to advance planned work, "code fests" bring together related communities to encourage interoperability, "hack-days" and "hack-weeks" teach hacking and making skills to diverse audiences without software backgrounds, like artists and scientists, "edit-a-thons" support intensive co-generation of encyclopedia content, and so on.

Taken together, these events offer new opportunities and challenges for cooperative work by affording explicit, predictable, time-bounded spaces for interdependent work and access to new audiences of collaborators. This one-day workshop brought together: researchers interested in the phenomenon, experienced event organizers, and participants interested in running their own events. The workshop aimed to facilitate consolidating existing research, sharing practical experiences, and understanding what benefits different event variations may offer, how they may be applied in other contexts, and how insights from studying these events may contribute to CSCW knowledge.

Workshop Organizing Committee	
Anna Filippova	Carnegie Mellon University
Brad Chapman	Harvard T.H. Chan School of Public Health
R. Stuart Geiger	Berkeley Institute for Data Science
James Herbsleb	Carnegie Mellon University
Arun Kalyanasundaram	Carnegie Mellon University
Aurelia Moser	Mozilla Science Lab
Arlin Stoltzfus	NIST
Erik Trainer	Carnegie Mellon University

Workshop Participants	
Areej Al-Wabil	Center for Complex Engineering at KACST and MIT
Dannon Baker	Johns Hopkins University
Daina Bouquin	Harvard-Smithsonian Center for Astrophysics
Ryan Curtin	Symantec Corporation
Dharma Dailey	University of Washington
James Damon	Harvard-Smithsonian Center for Astrophysics
Margaret Drouhard	University of Washington
Anna Filippova	Carnegie Mellon University
Brittany Fiore-Gartland	University of Washington
Eureka Foong	Northwestern University
R. Stuart Geiger	Berkeley Institute for Data Science
Liz Gerber	Northwestern University and Design for America
James Herbsleb	Carnegie Mellon University
Erin Hoffman	University of Washington
Jane Hoffswell	University of Washington
Chris Holdgraf	University of California, Berkeley
Krithika Jagannath	University of California, Irvine
Kenneth Joseph	Northeastern University
Arun Kalyanasundaram	Carnegie Mellon University
Javed Khan	Eindhoven University of Technology
Meris Mandernach	Ohio State University
Jonathan Morgan	Wikimedia
Aurelia Moser	Mozilla Science Lab

Arnab Nandi	Ohio State University
Brooke Sansosti	Carnegie Library of Pittsburgh
Steve Scallen	Microsoft
Brittany Smith	
Victoria Sosik	Google
Arlin Stoltzfus	NIST
Anissa Tanweer	University of Washington
Erik Trainer	Carnegie Mellon University
Jamie Whitacre	Project Jupyter

Introduction

Time-bounded collaborative events, sometimes called hackathons, data dives, code fests, hack days, sprints, edit-a-thons, map-a-thons, and so on, are exploding in popularity. In 2015, collegiate hackathons alone attracted over 54,000 participants across 150 events [7]. Conventional discourse is that ad-hoc teams of young coders compete in these several-day events, motivated to stay up all night by the appeal of free food, prizes and job offers. Yet there are variations in their design, purpose and adaptation across other fields and contexts that suggest they are a more broad form of cooperative work.

Such events may be non-competitive and oriented to specific themes and disciplines, like social good, to support participation from varying audiences [3]; they may differ on whether participants are present face-to-face or collaborate remotely, and the extent to which communication tools are used [8]; involve newly formed teams working on new projects or existing communities working on well-defined agendas [15, 11, 9]; they may be applied towards informal and collaborative learning [12, 5, 13, 10, 14], creating startups [4], innovative prototypes for arts and culture [2], civic open innovation [1] or strengthening interaction in specific scientific domains like computational biology [11, 9]. The hackathon model has even been applied to academic conference spaces through workshops exploring alternative models of creation, such as OCData @ CSCW ‘14 [6], and several events at other venues like CHI 2013-2014 [3].

However, to the best of our knowledge, work has yet to bring together these diverse threads of research and practice into a broader agenda. It is important to do so because despite these differences, hackathon-like events all share a common collaborative element: attendees team up with each other and use these spaces to ‘hack’ on new technologies and ideas, projects that are not within the scope of their regular work, or move forward work they otherwise would not be able to, due to either a lack of dedicated time or resources.

As such, these spaces introduce new and interesting opportunities and challenges for the study of Computer-Supported Cooperative Work. For example, hackathon-like events may provide unique opportunities for cooperation, by affording explicit and time-bounded spaces for individuals to work more interdependently, as well as providing access to new collaborators with needed background and experience, or existing collaborators who are otherwise difficult to reach, such as remote community members. Furthermore, some variations employed by distributed collaborative communities, such as regular sprints during yearly conferences, may support predictable interactions that can serve to strengthen existing social ties and develop new ones.

At the same time, working on projects that are outside of one’s normal workflow may provide challenges for continuity of this activity after the brief cooperative stint is over. For example, continuing projects in a virtual setting may require carrying over social and work artefacts that are not in easily editable formats and highly context dependent [15]. Time-bounded collaborative events may also provide different pressures on team dynamics during the event,

such as the need to go through team formation and development stages relatively quickly to be productive, as well as quickly dissipating dynamics and enthusiasm for completing projects when participants return to higher priority regular activities at the conclusion of the event.

This report presents an account of a one-day workshop organized to address some initial questions surrounding the hackathon phenomenon: What distinct practices do these events adopt, what goals and whose interests do they serve, and what makes them distinct from other cooperative activities? Do they achieve unanticipated benefits that can be leveraged in other contexts? How do we design more effective spaces to address specific event goals and what are the design tradeoffs? How can CSCW theory help in studying this space, and how do insights from this phenomenon add to CSCW knowledge? We invited both researchers and practitioners (including past event organizers, and individuals interested in running events in the future) to network, share ideas and have interesting conversations centered on the above questions, as well as to identify new areas of inquiry in this emerging space. All participants were asked to submit short papers describing their interest in the workshop, and relating their work and experiences to the above questions.

In the remainder of this report we describe the format of the workshop, including preparation and post-event activities. We summarize presentations and resulting conversations during the workshop, and highlight several themes that cross-cut the discussions. Finally, we present possible next steps for the community of researchers and practitioners working in this space.

Workshop

On Saturday, February 25, 2017 the “Hacking and Making at Time-Bounded Events: Current Trends and Next Steps in Research and Event Design” took place at the DoubleTree Hotel in Portland, Oregon. The workshop was held immediately preceding the annual ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW ‘17). It had several objectives:

- Facilitate networking between CSCW scholars and practitioners (both those who have experience putting on events and those who are curious about doing so);
- Develop an understanding of how to situate time-bounded events in the broader context of CSCW methods and theory;
- Identify and compile recommendations for organizers of events, as well as important tradeoffs; and
- Explore future directions for research in this area, including publication venues.

Format

Preparation

An 8-person organizing committee comprising a) CSCW researchers and b) practitioners with experience putting on events was selected with the aim to bring multiple perspectives to bear on event organization and advertisement, and participant recruitment and selection. CSCW researchers on the committee came from the Institute for Software Research at Carnegie Mellon University and the Berkeley Institute for Data Science at the University of California, Berkeley. Practitioners on the committee came from the Harvard T.H. Chan School of Public Health at Harvard University, the Mozilla Science Lab at the Mozilla Foundation, and the Institute for Bioscience and Biotechnical Research at the University of Maryland. We also sought feedback on the workshop design and proposal from researchers at University of Washington, Ohio State and University of Victoria working in this space.

The workshop was advertised through CSCW-themed communication channels, such as the conference's website, and the "Researchers of the Socio-Technical" Facebook Group. It was also advertised on a number of scientific domain-specific mailing lists such as `Open-Bio`. The organizers also leveraged their existing networks of contacts built from several years of doing work in this space, contacting researchers and practitioners from University of Washington, Ohio State University, University of Victoria, and developers of scientific software projects like MLPack, BioPerl, and Biopython.

All workshop applicants were asked to submit a 2-4 page paper describing their interest in one or more of the workshop themes, presented as a research idea or a story that drew from their own event experience. After all papers had been submitted, members of the organizing committee were randomly assigned submissions, and rated them on how well they represented the themes and their potential for discussion at the workshop. Each paper received at least two reviews.

Workshop Activities

A total of 33 participants representing 20 institutions attended the workshop. 57% of participants were female, and 43% male. After a brief introduction from the organizers, the workshop opened up with a short "boasters" session, in which each participant stood up and introduced themselves, their area of expertise, and what they hoped to take away from the workshop. This was followed by a series of four 90 minute blocks that comprised a combination of presentations, discussion, and additional activities designed to encourage the development of new collaborations.

There were a total of six 30 minute presentations from authors whose submissions several organizers agreed had the most potential to encourage significant workshop discussion related to the workshop objectives. The talks reflected a combination of work from researchers working in the space, practitioners with experience in organizing events, and participants interested in organizing future events. Each presenter had 15 minutes, balanced with a 15 minute discussion

that was a mixture of structured time led by a discussant and questions from the audience. Participants were invited to serve as discussants for the second half of each presentation block, based on their interest specified in submissions.

Our matching of presenters and discussants was purposeful, to enable cross-fertilization of ideas and diverse perspectives. We aimed to match participants and discussants by area of interest, such as organizing events in virtual spaces. However, we also aimed to vary the backgrounds of presenter and participant such that researchers were invited to discuss presentations given by practitioners, and practitioners those given by researchers. Discussants were asked to prepare by carefully reading the presenter's paper before the workshop to identify strengths, weaknesses, questions, and areas for further development. If the paper was theoretical, discussants were encouraged to consider practical implications, e.g., how and under what conditions its idea could be put in practice; if it was a practical account of an event, discussants were encouraged to consider how learnings could be abstracted towards our understanding of time-bounded events and opportunities for future research. The general structure for the discussion consisted of highlighting positive points, offering constructive criticism, and sharing preliminary thoughts and reactions to the paper.

The day also included a 60 minute "open space" session where volunteers proposed topics for discussion and others signed up to participate in the discussion. This session was meant to provide a loosely structured environment emphasizing the informal exchange of information and ideas. In the first few minutes, interested participants wrote down their ideas on sticky-notes. Following that, the organizers, with input from the audience, grouped the suggested topics into 4 high level categories. Four break-out groups formed corresponding to these categories.

Additionally, a 30 minute panel was held in order to think through event design, with an emphasis on opportunities and challenges for event success. The panel featured two workshop participants, one planning to run an event the following month, and the other interested in exploring the possibility of running an event, constituted the panel. One of the workshop organizers served as moderator.

The workshop concluded with a 30 minute session where the participants collectively identified issues not touched upon at the workshop, suggesting areas for future work and development.

Post-Event

In the week after the workshop, the event organizers sent a questionnaire to participants. The questionnaire included questions about participants' motivations for attending, their expectations, their satisfaction with different aspects of the event, as well as open-ended questions about what they liked about the workshop, and what they thought could be improved.

Results

A total of 19 papers were submitted to the workshop, all of which are available in the ppendix section of this report. Below we present summaries of presentations and discussions, and participants' feedback on the event itself.

Presentations

"A Typology of Hackathon Events"
Presenters: Meg Drouhard and Anissa Tanweer
Discussant: Victoria Sosik

Drawing on over 100 hours of ethnographic fieldwork, the presenters proposed an exploratory typology of time-bounded collaborative events. They described 3 types of events. *Communal* events are convened in order to develop capacities for a community: resources, infrastructure, practices, or culture. Participation is generally collaborative in nature. *Contributive* events aim to advance the aims of a larger pre-existing project, where the objective is to complete as much work as possible in a short amount of time. Finally, *catalytic* events are organized for the purpose of demonstrating the use of a dataset, or technology in a way that could spark new ideas and innovation.

This presentation raised multiple points for discussion, such as how helpful the typology could be for thinking about and designing events, particularly with respect to whether goals for different event types could be achieved simultaneously or whether they needed to be traded off, building community versus completing technical work. Questions were posed about how the typology considers team processes such as team formation and coordination between the teams. There was also interest in positioning the kinds of events discussed over the course of the day within the authors' proposed typology.

"Community and Code: Lessons from NESCent hackathons"
Presenter: Arlin Stoltzfus
Discussant: Dannon Baker

This presentation covered lessons learned from the organization and execution of over 10 years of hackathons conducted in the evolutionary biology community and sponsored by NESCent. The events were described as non-competitive, and lasting for four days or more. A key insight was the production of multiple outcomes, many of which are intangible, such as learning new skills or obtaining new experiences. The presenter communicated that many of the hackathon projects failed to continue after the events; instead the community itself seemed to grow, with instances of continued collaboration between attendees after the event had ended and interactions on newly-created mailing lists. These observations have implications for how to evaluate the success of these events and their longer-term impact.

Unsurprisingly, topics for discussion focused on practices for event organization, for instance in curating a diverse pool of participants. The point was raised that organizers tend to recruit

people like themselves, which could hinder diversity. The presenter recommended getting a diverse leadership team, and having the organizers do the recruiting. Another question was asked about what distinguishes some events from having successful continuity of projects, while others didn't. The presenter suspected that projects that aligned with participants 'day jobs was probably one of the most important factors.

The presenter also cautioned against remote participation at events, with the caveat that remote works for people who are already familiar enough with the project and with the people working on it. A point was then raised that having a remote organizing team can be helpful because it forces collaboration between remote and collocated participants from the inception of the event, smoothing interactions on the day of and post event.

"The CHI4Good Day of Service: What is Produced?"
Presenter: Liz Gerber
Discussant: Kenny Joseph

This presentation explored what is produced through philanthropic and civic hackathon events, drawing on interviews with participants, organizers, and representatives of nonprofit organizations. The authors pointed to multiple outcomes of these events: digital artifacts, expanded social networks, exposure to the design process, emotional changes throughout the day, and the shaping of interdisciplinary identities. Implications for the design of philanthropic hackathons were suggested, including offering design patterns and templates for participants, including opportunities for active reflection throughout the day, and connecting participants before and after the events.

One of the interesting points made in the discussion of this presentation was about sustainability, in particular the sustainability of organizers. There is a need to find ways to manage burnout that organizers may experience while at the same time getting people deeply engaged with the organization and execution of the event. Suggestions were made to understand the factors that influence turnover, such as compensation and motivations.

Another point was made that there is a lot of organizational theory that has yet to be brought to bear in order to more deeply understand these events, such as work around leadership, team roles, expertise and task matching.

"Building Something Amazing: 4 Years of Ohio State's Hackathon Program"
Presenters: Arnab Nandi and Meris Mandernach
Discussant: Erin Hoffman

This presentation described 4 years of competitive-style hackathons held at Ohio State University. The events aim to promote technical culture and promote technical talent, and have growth from 100+ participants in 2013 to 775+ in 2016. Basic tools for coordinating hackathon work such as GitHub, Slack, and Google Drive were introduced, as well as tools that did not work very well due to the additional overhead they imposed, such as Trello and Remind. The

presenters shared organizational strategies that seem to work well such as color-coordinated t-shirts for mentors. The presenters recommended creating and deploying post-event surveys to iteratively improve event execution. Examples included making sure teams were neither too small nor too large, and improving the food and work spaces provided. The presenters offered to make these surveys publicly available.

This presentation surfaced a contrast between exclusively student-run events and those run by the university. Events of the kind run by Ohio State are an example of bootstrapping hacking culture and the creation of a platform for informal learning. A challenge identified in the discussion was measuring the amount of learning that takes place. Participants agreed that it was necessary to go beyond self-report measures and GPA, and get education researchers involved in event evaluation to develop adequate metrics.

"Participatory Research in Open Science Events"
Presenter: Aurelia Moser
Discussant: Vassilis-Javed Khan

This presentation covered a range of hackathon-style events conducted over a 3 year period to promote open science on the internet. The events ranged from mostly collocated to mostly geographically dispersed. The talk provided insights into how to recruit a diverse group of participants (make sure the organizing team is adequately diverse), helpful technologies to support coordination and work activities (GitHub and spreadsheets), and practices for encouraging informal interactions among participants (assigning different badges to participants and offering prizes for finding others at the event with the same badge). The presentation also raised important issues about frustration and alienation at these events, particularly with respect to recruiting materials. The example given was an application form that asked if the attendee had a basic knowledge of Python. After clicking yes, a dialog box appeared with the text, "Are you sure you have a basic knowledge of Python?"

Conversations following this presentation centered on the tooling and resources Mozilla provides for event design. The presenter described a set of modules hosted on GitHub for event design, including assessment metrics and parameters such as types of activities, timelines for each, and thresholds for the number of event participants. Future event designers will be able to freely copy these modules and modify them as they see fit to meet their needs.

"Community Data Science Workshops"
Presenters: Jonathan Morgan and Dharma Dailey
Discussant: Brittany Fiore-Gartland

The presenters spoke to the execution of 4 years of data science workshops, learning-based hackathons for "total beginners" to data science and programming, at the University of Washington. The presentation emphasized striving to be as accessible and inclusive as possible, creating problem solvers not programmers. This included practices like keeping technical jargon to a bare minimum (i.e., deliberately avoiding terms like "recursion, class, unit

testing," and so on), not using the term "hack" anywhere in their workshop materials, and being tool/platform agnostic for people who already know other tools.

Audience members were curious about the role of the physical event space in achieving the goals of learning and building confidence with the tools and techniques introduced. The presenters favored a classroom style space with learners arranged in rows and the instructor at the front of the room. This makes it easy for mentors to access participants. The style of interactive programming used, where an instructor is coding live and participants can see mistakes being made was thought to be helpful for newcomers who may be intimidated by their abilities when compared to those of experts.

Open Space

"Tools"

The first breakout group focused on the practical aspects of event organization, such as the different stages of event organization and lessons in each stage, in particular, the kind of tools organizers have employed at different stages.

In the first stage, identified as the "Preview" stage, organizers aim to get the word out about the event through as many different forms of media as possible: listservs, flyers, postcards, facebook, chalking on campus, social media, and individually reaching out to get new participants.

In the next "Signup" stage directly preceding an event, more targeted promotional activity happens featuring classroom visits (for collegiate and school events), promotion of sign-up pages for the event, encouraging team sign-up, and communicating the rationale for teams to sign up.

Following this, in the "Team Building" stage, organizers noted employing strategies such as gamifying the process of team formation, and using visible markers to help participants find others with relevant skill sets (such as varying stickers for different skill sets). Notably, the team building stage can begin several weeks before an event, or occur only on the day of the event, depending on the event organization. One organizer of a large hackathon at Microsoft described an online tool called Hackbox, that the organizers have developed to support finding teammates at very large events. The tool allows participants to set up individual profiles far in advance of an event, describing details like what they working on, what skills they have that others would find valuable, and their personal passions at work. The tool also allows to register project ideas before the event as well, select project categories (such as, who is this project for or if it is part of a special challenge at the event), and list specific skills the project needs. The tool allows to search for people or for projects, and can look at a given user profile and perform an affinity match to projects for best fit.

An additional pre-event stage was identified as "Getting Ready" -- a stage that uses pre-events to support higher levels of commitment from participants such as holding meetups among registered participants, offering training on how to make pitches, gathering external partners and

facilitating meetings between event partners (such as non-profits for whom hackathon solutions may be developed), and conducting workshops to learn about tools that the project team may need to work with. One important element to note about pre-events is that if not everyone in the team is able to participate, the event has to produce something that is shareable to the rest of the group who can't physically make it there.

A post-event stage that the organizers termed "Showing off" could involve a variety of tools for presenting the final product -- a traditional presentation, is common although some events use a video demo as a final product. Encouraging participants to create a video has proven to be valuable to organizers because it creates a repository of resources. An additional means of showing off projects is a "science fair" style approach, reminiscent of demo and poster sessions at conferences, allowing both event participants and those who didn't participate in the hackathon to walk around and explore and interact with the final products.

The group also discussed the final stage of "Momentum" or sustainability, that is, how to keep people excited throughout the year about their projects and future events.

"Diversity, Remote Participants, Newbies"
This group discussed 3 topics: remote participation, welcoming newbies, and diversity.

Several levels of remote participation were discussed along with their issues. One kind of remote participation was termed the satellite hackathon, where the whole team is collocated with each other, but geographically dispersed from all other teams. In this case it may be helpful to have a wall of monitors to show what all other teams are working on, and to remind people that there is more going on they might want to be aware of. Another kind of dispersion, is where you may have isolates who are geographically dispersed from their team. In this case it helps if the remote participant closes their office door and puts up a sign so they do not experience distractions from others within close physical proximity. This kind of dispersion is perhaps more problematic since the participant does not get the benefits that come along with being f2f, such as seeing what people are looking at, overhear conversations that spark impromptu meetings, and recognize when people are puzzled or deep in thought. Having the remote person buddy up with someone who is face to face may alleviate some of these issues, but will likely not scale up.

Under the topic of welcoming newbies, the issue of frustration was raised and how to go about identifying and managing it. Frustration among women was highlighted, with one group member pointing out that women consistently report lower self-efficacy with technology compared with men. One participant in the group told a story about how at an event, several women experienced errors in the software they developed. They thought they had introduced the errors themselves, when in reality there was a bug in the code that they were re-using. Another story was brought up about a woman feeling so frustrated at an event that she wanted to give up and go home. Suggestions for managing frustration were given, such as approaching potential problems from the side, i.e., never asking directly if there is a problem, rather asking things like "How's it going?" or "What do you think of the food?" It was suggested that ways of dealing with

frustration should be linked to the purpose of the event. For instance, in learning-based hackathons it may be useful for instructors to ask questions like, “What do you mean by this import statement?”

Under diversity, the topic of recruiting was raised. Regarding recruiting, the issue of how language influences who shows up was explored. The term *hackathon* was recognized by the group to imply competition, which can be a turn on for some folks, and a turn off for others. *Jam* was suggested as a neutral alternative, but even then members of the group noted that it suggests “if you show up, you better be ready to jam.” The group also recognized that the identity of the event has to be formed beyond the name, through informal communication channels used by organizers and participants, to recruitment materials, to codes of conduct. In other words, the name can’t do all the work. Language and messaging are where people determine, “Yes this is for me” or “No, this is not for me.”

“Hackathons as Representations of the Changing Nature of Work”
This group discussed hackathons as reflecting the changing nature of work. The characteristics of this shift in work include collocation, remote teams, modularity, and providing a space separate from work, where people can do things they would not normally do. In this kind of work, traditional work processes may not necessarily apply and intentions may vary. Traditional metrics such as productivity and continuation of projects may be shifted. In other words, there may be different goals by which success is measured.

Topics not discussed yet at the workshop were explored, such as boundary conditions (e.g., which kinds of projects might hackathons not be appropriate for), who the stakeholders are in the community, shared understanding, grounded communication, and conflicts in traditional in work priorities such as failing fast but sustaining relationships. For instance, one criterion for an event to be considered a hackathon is if the goal isn’t to finish a project, but rather to set something up that can be built upon - one possibility would be as a launch pad for ideas or work that can be funneled into an academic semester/quarter.

“Measuring Outcomes of Hackathons, Specifically Educational Outcomes”
The big question discussed in this group was how to show the impact of hackathon-style events. The main challenge identified is that with respect to assessment, one size doesn’t fit all because of the diversity in event objectives and design variations.

Two approaches to evaluation were discussed. One is to survey hackathon participants, but it is difficult to do well and meaningful conclusions cannot be drawn from small events. The second approach is follow-up interviews, but this does not scale well beyond a dozen participants.

An alternative way of measuring outcomes was to do sort of a meta-hackathon with a small representative set of people who had participated in one or more of the events. Group members identified a sub-literature in engineering education as a possible resource for finding examples in how to obtain value from a subset of people reflecting on their experiences.

Measuring the long-term impact of time-bounded events was also discussed, where members of the group suggested tracking event participants and their participation in their social groups over time.

Panel: Thinking Through Event Design: Challenges and Opportunities for the Future

The panel opened up with short introductions from the two panelists and stories of how they got to the point of organizing, and thinking about potentially organizing, events.

Panelist #1, a lead developer of a Machine Learning library, talked about his experience leading the library's participation in the distributed mentoring program Google Summer of Code (GSoC). One of the problems he has faced is identifying skilled potential contributors; the library receives many applicants each year (119 in the most recent year), and only about 10 are suitable. This is overwhelming for small communities like his. Creating an issue for beginners to solve is a large time commitment, and is often something he could do much more quickly himself.

As a consequence, he was skeptical about being able to find skilled participants for a hackathon-style event. Compounding the problem, he was unaware of people in research labs and other institutions using the software who might be interested in attending.

To help filter out unqualified candidates to onboard, it was suggested to impose a badge requirement for applicants, e.g., a prerequisite of having X number of pull requests accepted. A related suggestion was to have themes for pull requests to help ensure the submitted changes address issues deemed important and desirable by the core team.

Panelist #2, a fellow at the Berkeley Institute for Data Science, talked about an upcoming hackathon-style event that he organized, focused on encouraging people to document their open-source software projects rather than adding new features. One of the concerns he raised was thinking about how to measure success of the event, admitting that metrics like number of overall documentation-related commits made during the event were useful but imperfect. Quality of the documentation is important as well.

To encourage quality documentation, before the event panelist #2 made a collection of resources available to projects seeking new contributors and contributors themselves. These resources included checklists for encouraging contributions, for instance making it clear how documentation is organized and having examples of good forms of documentation. They also included checklists for new contributors, which had items like improving the language used in examples and looking for issues labeled as documentation.

He stressed the use of tools for the event, as it would be remote rather than face-to-face. GitHub would be used as a platform for the event because it allowed people to engage in discussions about what they would be working on during the event, and track all participants'

work. He also emphasized using Slack to facilitate real-time communications between remote participants.

Questionnaire Feedback

Overall, questionnaire respondents were motivated to learn new things and build social connections with other participants.

All respondents indicated that they either somewhat, or to a very large extent met these goals. In addition, over half of the respondents (64%) meeting 2 or more people who they could pursue new projects within the future.

The vast majority of respondents (86%) said they felt the objectives of the workshop were clear (marked by a 4 or 5 on a 5-point scale from strongly disagree to strongly agree). The same percentage of participants said they were satisfied with the quality of output from the workshop, that their contribution was reflected in that output, and that it was feasible to meet the goals of the workshop in the time allotted.

The most prominent theme in what people liked about the event was the variety of content and perspectives of workshop participants:

“It brought together a diverse group of thinkers to work on an increasingly important topic within the broader research community”

“I got to meet so many people with different perspectives”

“Mix of attendees was excellent”

In addition, the workshop exposed participants to new things:

“As a novice, I was introduced to the theory and practice of time-bound events. I feel confident I can take this information and apply it to any time-bound events I am part of in the future.”

When asked about what could be improved about the event, participants indicated that the workshop was too ambitious in its goals:

“We really needed more time to make [the event objectives] happen”

“The next event would be more successful if it focused on something narrower.”

Discussion

Themes

Cross cutting the presentations, discussions, and open space sessions were several key themes:

Affect

The theme of affect came up in Gerber's presentation, Morgan's presentation, as well as the open space session on newbies, diversity, and remote participation.

Emotion is central to the hackathon experience, be it the thrill of working on something new or disruptive, or the frustration that comes with not getting something to work. How can we design events that help avoid or at the very least manage, negative emotional responses and support positive ones? What practices should organizers and leaders follow? Morgan gave an example of live-coding by instructors during learning-based hackathons, where mistakes make learners feel less intimidated. How might these practices need to be adapted for other flavors of events, e.g, those geared primarily toward advancing technical work?How does the language used in hackathon recruiting materials, codes of conduct, agendas, and organizer discourse influence participant affect? Coming up with tools and techniques for measuring affect will be of great help in answering these questions as well as recognizing emotion as a tangible rather than intangible output of hackathon-style events.

What are we sustaining?

The question of what is being sustained with a hackathon was an issue raised in several presentations at the workshop.

Drouhard's presentation identified three possibilities: the community, hackathon projects themselves, and ideas. In his description of bioinformatics hackathons, Stoltzfus identified projects and the community, but cautioned that sustaining projects is difficult unless there is an obvious link to the participants' day jobs. Sustainability of the organizers and leadership was highlighted in Gerber's presentation. How do communities manage burnout and turnover while simultaneously keeping participants deeply engaged?

Language matters

The topic of language, and its role in shaping the hackathon experience was a subject brought up in various presentations and informal conversation. Gerber's presentation posed questions about how the language used both in how the event is named and language used by participants and organizers, affects participation. For instance, phrases like "fail early fail fast" may work to motivate those proven to build market-changing products, and who can financially afford it; to novices however, "failure" may have much more dire consequences that are not worth risking in a hackathon setting.

The language used should be appropriate for both the aims of the event and the desired participant pool. The language above is not likely to be appropriate for instance, for an event aimed at learning. For these types of events, using the minimum amount of technical terms, as discussed in Morgan's presentation, may prove to be helpful. Further, attention should be given to how the event is named. Terms ike "hack" and "marathon" may especially be a turnoff for newcomers. It is also important to be mindful about language in event materials like survey instruments and CFPs. Workshop attendees in the diversity, remote participants and newcomers discussion group noted the importance of having event roles other than "programmer." Moser's talk called attention to an online survey that repeatedly questioned participants' basic programming language. This caused some participants to feel uncomfortable and doubt their ability to participate.

Next Steps

At the conclusion of the workshop, participants agreed that several topics needed more attention going forward:

- Measurement

There are a number of desired outcomes of these events, such as learning, community building, artifacts produced, innovation, and so on, but we lack good ways of measuring them. Educational, learning based events typically rely on self-reported measures of learning like GPA, which can vary due to a number of factors, including class difficulty and strictness of instructors and cannot be confidently compared across institutions. Once we have ways of measuring event outcomes, we need to be able to connect them with objectives at multiple levels of perspective, e.g., from the participant, organizer, and organization. How does the outcome from a corporate hackathon-style event contribute to broader organizational goals and objectives? Evidence from our own work suggests that when the motivations of participants differ significantly with organizer goals, it is associated with less successful event outcomes. If, for example, a large pool of newcomers attend who are motivated to learn new tools and network with other participants shows up to an event aimed at producing prototypes and innovative software. These goals may conflict, as inter-group communication and onboarding will potentially take a lot of time away from actually getting work done. Additionally, it is important to consider measures that encourage accountability, that is, are connected to the objectives of the community or organization running the event.

- Diversity

Designing for diversity on multiple dimensions (e.g., age, race, gender, role, expertise) is a topic worthy of much more attention. As we learned over the course of the day, events that don't design for diversity have the potential to alienate potential participants or increase the likelihood for frustration and other negative emotional responses.

Several strategies were identified for supporting better diversity. One of the most important strategies identified was to ensure the organizing committee is diverse, as such a group would

be more sensitive to issues of diversity in event organization. Additionally, de-emphasizing the competitive nature of an event, or removing competition altogether was identified as being helpful in attracting a more diverse range of participants. Sometimes, however, competition is necessary for the event design, and an alternative strategy may be to focus the event on social good themes, and/or provide prizes that are donations to charities rather than a chance to win physical goods. Furthermore, using language that avoids using the word “hack” in favor of “workshop”, and minimizing the use of language about “code” if more than software outcomes are desired were identified as helpful strategies to attract both functional and demographic diversity of participants.

It is also important to encourage group processes that support diverse teamwork: previous work finds that participants who use principles from brainstorming (that is withholding criticism, ensuring everyone is able to contribute ideas, and building on all ideas generated rather than discarding them) is associated with greater team satisfaction, particularly for minority members of diverse teams [16]. Having a well designed and accessible code of conduct can also serve to set expectations about interactions within diverse groups. Finally, as discussion during the workshop suggested, engaging in pre-event activities that facilitate interaction may help in developing a shared language that smoothes coordination and helps build trust needed to work openly with each other.

- Future communication and events

At the workshop’s conclusion, participants agreed to use the `Hackathon-workshop`[1] mailing list to continue discussions around these topics, share resources (e.g., articles, slide decks, event planning materials), and facilitate collaborations. Its purpose is to maintain, and possibly grow the community of researchers and practitioners working in this space. As mentioned previously, post-questionnaire feedback revealed that the day’s agenda was too ambitious. One of many follow up possibilities would be to use the mailing list to plan a follow up workshop, centered on the themes of measurement and diversity. An additional element identified was the need to reach out to groups not yet represented at this workshop: researchers with a strong educational background to support more detailed discussion about facilitating learning outcomes, as well as groups that are organizing such events at scale, such as Major League Hacking (MLH), a non-profit that supports the organization of over 200 collegiate events a year.

Conclusion

Time-bounded collaborative events like hackathons, codefests, sprints, edit-a-thons, and data dives are seemingly everywhere—from college campuses to city libraries to firms. As such they have different audiences, with different motivations and intended outcomes. There are many other parameters—from the way teams are formed, to the communication tools used, to the geographic distribution of team members—that make each event different in its own way. How to configure these parameters to achieve intended outcomes is critical to event success.

[1] https://lists.andrew.cmu.edu/mailman/listinfo/hackathon-workshop

The purpose of this 1-day workshop on time-bounded events was to gather researchers and practitioners together to network and share ideas about how to design effective spaces to address specific event goals, the tradeoffs involved, and how theory can help in studying the phenomenon. The presentation/discussant format seemed to facilitate the cross-fertilization of ideas between both groups of participants, sparking new conversations about these events' relationship with sustainability, recognition of intangible outcomes like emotion and learning, and the role of language in shaping hackathon experiences. Overall participants had positive impressions of the workshop, saying it was effective in helping them learn something new and build social connections with other attendees. At the same time, some participants felt the day's agenda was too ambitious, suggesting that more time was needed to address key topics like measurement and diversity in sufficient detail. One of many follow-up possibilities would be to organize a future workshop that is less broad in scope, focusing specifically on one of these topics. The mailing list set up after the event can be used to plan such collaborations. Our hope is that the community of researchers and practitioners working in this space continues to collaborate, share resources, and grow.

References

1. Esteve Almirall, Melissa Lee, and Ann Majchrzak. 2014. Open innovation requires integrated competition-community ecosystems: Lessons learned from civic open innovation. *Business Horizons* 57, 3 (2014), 391 – 400. DOI: http://dx.doi.org/10.1016/j.bushor.2013.12.009
2. Gerard Briscoe. 2014. Digital innovation: The hackathon phenomenon. *Working Papers of The Sustainable Society Network* 2014 (2014).
3. Adrienne Decker, Kurt Eiselt, and Kimberly Voll. 2015. Understanding and improving the culture of hackathons: Think global hack local. In *Frontiers in Education Conference (FIE)*, 2015. 32614 2015. IEEE. IEEE, 1–8.
4. Marthe Dehli. Hackathons as a ground for creating start-ups: Evidence from THE Port 2014. Ph.D. Dissertation. CERN.
5. Allan Fowler. 2016. Informal STEM Learning in Game Jams, Hackathons and Game Creation Events. In *Proceedings of the International Conference on Game Jams, Hackathons, and Game Creation Events* (GJH&GC '16). ACM, New York, NY, USA, 38–41. DOI:http://dx.doi.org/10.1145/2897167.2897179
6. Sean Goggins, Andrea Wiggins, Susan Winter, and Brian Butler. 2014. OCData Hackathon @ CSCW 2014: Online Communities Data Hackathon. In *Proceedings of the Companion Publication of the 17th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW Companion '14)*. ACM, New York, NY, USA, 317–318. DOI: http://dx.doi.org/10.1145/2556420.2558865
7. Major League Hacking. 2016. Hackcon IV: State of the League. (2016). https://www.youtube.com/watch?v=NeH7RDEr4Jw

8. Mozilla Science Lab. 2016. Mozilla Science Lab Global Sprint 2016. (2016). https://science.mozilla.org/ programs/events/global-sprint-2016
9. Hilmar Lapp, Sendu Bala, James P Balhoff, Amy Bouck, Naohisa Goto, Mark Holder, Richard Holland, Alisha Holloway, Toshiaki Katayama, Paul O Lewis, and others. 2007. The 2006 NESCent phyloinformatics hackathon: a field report. *Evolutionary Bioinformatics Online* 3 (2007), 287.
10. Miguel Lara and Kate Lockwood. 2016. Hackathons as Community-Based Learning: a Case Study. *TechTrends* 60, 5 (2016), 486–495. DOI: http://dx.doi.org/10.1007/s11528-016-0101-0
11. Steffen Möller, Enis Afgan, Michael Banck, Raoul JP Bonnal, Timothy Booth, John Chilton, Peter JA Cock, Markus Gumbel, Nomi Harris, Richard Holland, Matúš Kalaš, László Kaján, Eri Kibukawa, David R. Powel, Pjotr Prins, Jacqueline Quinn, Olivier Sallou, Francesco Strozzi, Torsten Seemann, Clare Sloggett, Stian Soiland-Reyes, William Spooner, Sascha Steinbiss, Andreas Tille, Anthony J. Travis, Roman Valls Guimera, Toshiaki Katayama, and Brad A. Chapman. 2014. Community-driven development for computational biology at Sprints, Hackathons and Codefests. *BMC Bioinformatics* 15, 14 (2014), S7. DOI: http://dx.doi.org/10.1186/1471-2105-15-S14-S7
12. David Munro. 2015. Hosting hackathons a tool in retaining students with beneficial side effects. Journal of Computing Sciences in Colleges 30, 5 (2015), 46–51.
13. Arnab Nandi and Meris Mandernach. 2016. Hackathons As an Informal Learning Platform. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education (SIGCSE '16)*. ACM, New York, NY, USA, 346–351. DOI: http://dx.doi.org/10.1145/2839509.2844590
14. Anders Sigfridsson, Gabriela Avram, Anne Sheehan, and Daniel K Sullivan. 2007. Sprint-driven development: working, learning and the process of enculturation in the PyPy community. In *IFIP International Conference on Open Source Systems*. Springer, 133–146.
15. Erik H. Trainer, Arun Kalyanasundaram, Chalalai Chaihirunkarn, and James D. Herbsleb. 2016. How to Hackathon: Socio-technical Tradeoffs in Brief, Intensive Collocation. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16)*. ACM, New York, NY, USA, 1118–1130. DOI: http://dx.doi.org/10.1145/2818048.2819946
16. Anna Filippova, Erik Trainer, and James D. Herbsleb. 2017. From diversity by numbers to diversity as process: supporting inclusiveness in software development teams with brainstorming. In *Proceedings of the 39th International Conference on Software Engineering (ICSE '17)*. IEEE Press, Piscataway, NJ, USA, 152-163. DOI: https://doi.org/10.1109/ICSE.2017.22

Appendix

Aurelia Moser
Community Lead
Mozilla Science Lab - Mozilla Foundation
Proposal for Hackathon Workshop 2017

Participatory Research in Open Science Events

Since its 2013 inception, the Mozilla Science Lab has worked to build community around open science, building capacity to support community programs and developing prototypes and projects that foster scientific research on the open web. In this effort, we organize annual “tentpole” events like our Global Sprint and Mozfest to celebrate the scientific contributions to the open internet and open source projects worldwide.

Here is a short-list of some of the programs and projects we run:
Events: https://science.mozilla.org/programs/events
Fellowships: https://science.mozilla.org/programs/fellowships
Study Groups: https://science.mozilla.org/programs/studygroups
Projects: https://science.mozilla.org/projects
People at Mozilla Science: https://science.mozilla.org/people

This past year, we continued with our second cohort of a Science Fellows program to support post-doctoral candidates piloting research and experimentation with goals to education, sharing, and connection with scientific communities. We've also run global and local scale “sprints” to onboard our community members to the ethos of open science and developed curriculum and training materials to inform our events with best practice in data science, peer review, and open source research. Part of our plan for the future of the open web involves partnering with like-minded groups who share in our enthusiasm for open research, and strive to broaden the community receptive to open science.

To these collaborations, we bring a strong community of supporters, the dedicated time and attention of our staff, and the tested curriculum in open science practice and planning that forms the foundation of our programs to date. Our organization coordinates additional ad-hoc events each year, and partners with other “open” focused organizations and institutions to run and deploy these events, the following sections will review the types of hackathon and participatory research events we support each year, organized from highest-involvement to moderate endorsement. Each event implies a different approach to preparations, designs for a different timeline and resource dedication level, and in many cases, appeals to a slightly different audience.

***

Tentpole Events

Global Sprint

Each year, roughly at the beginning of June, the Mozilla Science Lab organizes a global hackathon or sprint at local “sprint sites” in over 30 locations globally for over 30 hours of continuous hacking across all timezones. We build mini-event websites for all participants, and curate a strong list of open source projects in the sciences that are available for remote contribution from any hackathon participants around a few key themes: open data, open educational resources, citizen science, and tools. Likewise, we offer mentorship for project leads, coordinate regular check-ins with regional coordinators and site participants over Vidyo chat, and host conversational text-chats via Gitter channels according to those regional blocks. We collect metrics on event participants and use this as an opportunity to engage more community in our programs.

For the sprint, our start time is several months in advance for prep, our resourcing is between 2 and 6 people consistently throughout this timespan, and our audience is global/mutli-lingual.

MSL Event Site: https://science.mozilla.org/programs/events/global-sprint-2016
Projects: https://science.mozilla.org/programs/events/global-sprint-2016/projects
Example Site: https://ti.to/mozilla-science/gs2016-newyork

Press:
Site in Utrecht
Site in Madrid
EPCC Promo Post

Mozfest

In the Fall, the Mozilla Foundation hosts a festival for the open internet in London, UK. The event is hosted at Ravensbourne college, and includes an opening night Science Fair for project demos and sharing resources, and it’s followed by two days of sessions on 9 floors in the college. The Open Science floor is usually floor 9, and we coordinate workshops and sessions based on community proposals, hosted over the remaining two days of the festival.

For Mozfest, our start time is a year in advance for prep, our resourcing is an event planning team of 2 and an entire production (6 designers and devs) team for prep, plus the entire Mozilla Foundation for proposals and floor preparation consistently throughout this timespan, and our audience is global/mutli-lingual.

Event Site: https://mozillafestival.org/
Schedule: https://app.mozillafestival.org/#_space-open-science

Working Open Workshop (WOW)

In February 2016, the Mozilla Science Lab launched its first Working Open Workshop in Berlin, a series of presentation sessions and workshop handouts to onboard scientist participants to open research, the protocols of version control, open access licensing, and data management as well as helpful tips for designing events and contributor roadmaps and codes of conduct to ensure that all community members feel welcome and informed about their open source projects or programs. This workshop will be redeployed over the next few years in 10 other communities/cities with slight modifications on the same theme.

For the WOW, our start time was 2 months in advance for prep, our resourcing is between 2 and 6 people consistently throughout this timespan, and our audience is selective and usually limited to a carefully chosen group of at-max 30 mentees.

Event Site: https://science.mozilla.org/programs/events/working-open-workshop-february-2016
Presentations, Handouts, Transcripts: https://github.com/mozillascience/working-open-workshop
Activity Site: http://mozillascience.github.io/working-open-workshop/

Space Apps + Science Hack Day

Last year, Mozilla Science hosted the Brooklyn/NY Space Apps Hackathon, a collaboration with NASA who issues challenges annually and allows locations all over the world to host hackathon sites to tackle those challenges. This year, one of our teams made it to final judging in the origami category, with a project to develop self-folding repair robots for rovers on Mars.

For SpaceApps, our start time was 1 month in advance for prep, our resourcing was one person throughout this timespan, and our audience was local to the New York and Brooklyn area.

Site: https://2016.spaceappschallenge.org/locations/brooklyn-ny-usa
Our Science Hack Day site: http://sciencehackday.nyc/
Pre-hackathon Datacamp: http://sciencehackdayny.github.io/data-camp-16/
Github with Projects: https://github.com/sciencehackdayny

Conferences + Science Fairs

GET Conference

Our 2015 Mozilla Science Lab Fellow, Jason Bobe was one of the coordinators of the Genomes Environments and Traits conference, which featured a day of talks related to open and participatory biomedical research as well as a science-fair-style “labs” event where attendees

could walk from booth to booth, signing up for biomedical research studies, many with mobile applications for submitting samples and user data.

For GET, our start a few weeks in advance for prep, our resourcing was limited since most of the coordination went through our fellow, Jason, and our audience was mostly biomedical researchers and participatory medicine scholars in the Boston area (it was hosted at Harvard Medical school)..

Event site: http://www.getconference.org/
GET Labs: http://www.getconference.org/get2016/labs.html
Open Humans: https://www.openhumans.org/

GeoJourNews

Our 2015 Mozilla Science Lab Fellow, Joey K. Lee was a featured speaker at a conference we co-organized and sponsored called GeoJourNews, a day of talks about geospatial science, journalism, and storytelling with cartography.

For the conference, our start time was 4 months in advance for prep, our resourcing was just myself and collaborators at CartoDB consistently throughout this timespan, and our audience was a mixture of journalists and geospatial scientists.

Event site: https://nvite.com/GeoJourNews/ab4d

Code for America Summit + Mozfest Review

Post-Mozfest, we sometimes organize small meetups like this one, to review the content of the sessions for those who could not attend, and to connect our event with other events, like the CFA summit.

For the CFA Summit / Mozfest meetup, our start time was less-than a week in advance for prep, our resourcing was <1 person throughout this timespan, and our audience is was New York-based and civic-tech focused.

Event site:
https://www.eventbrite.com/e/mozfest-cfasummit-recap-by-betanyc-microsoft-civic-hall-mozilla-tickets-29257946305#

***

As research becomes increasingly computationally and data intensive, the need for collaborations and co-designed event programming supported by a broad network of organizations also increases. We look forward to that collaboration, and are always on the lookout for more opportunities to partner and share.

Building Something Amazing: 4 years of Ohio State's Hackathon Program

CSCW Application

Your name, title, affiliation, and email.

Arnab Nandi, Assistant Professor, Computer Science & Engineering, arnab@cse.osu.edu (Faculty Director, OHI/O Program)

Meris Mandernach, Associate Professor, University Libraries, mandernach.1@osu.edu (Faculty Director, OHI/O Program)

Title, Abstract, 3 Keywords

Building Something Amazing: 4 years of Ohio State's Hackathon Program

To foster a tech culture at Ohio State and cultivate technical talent in Columbus and the Ohio region, "HackOHI/O" — Ohio State's Fourth Annual 24-hour Hackathon and programming contest was held on the weekend of November 19th-November 20th. The event has grown substantially from 100 to 200 to 500 to over 775 undergraduate and graduate student programmers. HackOHI/O 2016 also included participants from 26 universities who built projects, competed for prizes, and were judged by faculty on categories including technical difficulty, creativity, usefulness, and presentation. In its 4th year, HackOHI/O has become one of the 25 "Signature Events" on campus. Each year participants are surveyed to ensure quality improvements year over year. We would share our experiences on program design principles, implementation details, quantitative metrics, managing rapid growth, impact on culture on campus.

Keywords: Informal learning, student culture, event design

A description of one or more themes of particular interest to you that are related to the workshop topic. This may be presented as: an extended abstract summarizing a research idea, a recounting of an experience with a related event, or a story that draws from your own research or event experience.

Design variations: We would like to share our experiences from 4 years of running the "OHI/O" informal learning platform. This program focused on fostering a tech culture at Ohio State University includes an annual hackathon (HackOHI/O), a hardware focused hackathon (MakeOHI/O), grants to work on projects and showcase their projects (ShowOHI/O), and travel funds to attend other hackathons. This would involve discussions around aspects of planning and implementing both a hardware and software focused hackathon, the changes in planning teams, and growth of the program as well as the planning related to the program.

Over the past four years of HackOHI/O the event has grown from 100 to 200 to 500 to 775+ participants. Additionally, the planning team has grown from 3 to 6 to 20 to 30 student and staff planners. In 2014 the program added a hardware hackathon (MakeOHI/O) which has grown from 50 to 100 participants. In 2015 discussions started to include ShowOHI/O an event where students could showcase their continued work of projects from events. Participant quotes from 2016 event:

> *“I came in with no computer, no team, and no ideas. All of those things were eventually taken care of and I left with a working android app. What more could I ask for?”*

> *“I'm definitely applying for next year's. This was my first hackathon, and it showed me how fun it was to code outside of classes.”*

> *“I really enjoyed diving into uncharted territory and working so hard to make something happen. It was very fun to see people having success and make something myself.”*

> *“Having an environment where I'm surrounded by people who are coding just like me. I felt much more motivated to learn.”*

The event now has spinoffs and satellite events, such as the Hardware-focused “MakeOHI/O” which is its own annual event, the Ohio State Wikipedia club's editing sprints, and 10+ pre-events ramping up to the Fall Hackathon. HackOHI/O is now a “Signature Event” at OSU, alongside homecoming weekend etc as one of the 25 top student events on campus. This is a grassroots movement 4 years in the making — we were able to assemble a team of 3, then 7 and now 25+ students, staff and faculty members to move the needle of tech all across campus.

Practical support for event organizers: Over the course of the past four years, the team has developed a number of streamlined approaches for planning a growing event. From the past four years the team has surveyed participants about the event for improving the experience. Additionally, during the 2016 event judges and mentors were surveyed about their experiences. We will share ideas related to registration, judging, general atmosphere, and regular team meetings. We will also share budgets, sponsorship goals, and best practices for planning such events.

A short summary of your background, interest in this area, and motivations for participating in the workshop.

Background of OHI/O & Interest in this Area:

We are co-directors of "OHI/O", a platform for informal learning for students in tech. We noticed that there was a culture gap in technology — we wanted students to be excited and self-motivated about building new technology beyond grades and graduating. Thus, we started an annual hackathon event in 2013, which has doubled in size each year. In 2016, there were over 775+ participants (130+ women), 26 universities, 30+ sponsors, and 300+ mentors, judges, showcase attendees this year.

Such events foster a culture of learning and building beyond just the classroom, where students learn at the event. We have noticed that our multi-year hackathon platform provides an extremely effective environment for informal learning and culture change, without a top-town / formal instructional climate for students.

Motivation with the Workshop:

1. We would like to share 4-years of our experiences of designing, building, and running a University-level hackathon program from scratch to the largest in the state through grassroots collaboration with students and other staff members.
2. We would like to share our insights into "patterns of work" from analysis of participants github repositories that show, through commit logs, different styles of team work during the same hackathon event.
3. We would like to discuss with peers and learn more about how to design / instrument our hackathons better for informal learning and collaborative work.

Backgrounds of the applicants:

Arnab Nandi is an Assistant Professor in the Computer Science and Engineering department at The Ohio State University. Arnab's research is in the area of database systems, focusing on exploiting user behavior to address challenges in large-scale data analytics and interactive query interfaces. This involves solving problems that span the areas of databases, interactive visualization, human-computer interaction, and information retrieval. He is 2016's recipient of IEEE TCDE Early Career Award for his contributions towards user-focused data interaction. Arnab is also a recipient of the US NSF CAREER award, a Google Faculty Award, and a Yahoo! PhD Fellowship. Prior to joining Ohio State, Arnab received his PhD from the University of Michigan, Ann Arbor in 2011.

Meris Mandernach is an Associate Professor in the University Libraries at The Ohio State University. She is responsible for developing, refining, assessing and sustaining an evolving program of research services for faculty, researchers and students throughout the university. She led a team that developed the services and space for the new Research Commons at OSU's library. Her research focuses on interactive library exhibits, planning and implementing interdisciplinary research services and spaces, and understanding user needs to improve library services. She has over 10 years of experience in various academic libraries in both public services and technical services and has previously worked at James Madison University and Loyola University Chicago.

If relevant, you may provide links to additional online materials in the PDF.

2016 Report: http://go.osu.edu/hackohio2016report
This is a comprehensive 9-page report for the 2016 HackOHI/O hackathon, which includes demographics, pictures, media highlights, sponsor details, winning teams, quotes / tweets from leadership / faculty / students, ratings from our post-event survey, and more.

Progress of 2013--2015:
ACM SIGCSE `16 Paper, ***"Hackathons as an Informal Learning Platform"***:
http://arnab.org/files/hackathon_informal_learning.pdf

Startups & Tech that was "Born at OHI/O":
https://u.osu.edu/technologycommercialization/2016/05/03/track-my-receipt-to-the-app-store/

Music Videos:
Each year, we have a tradition of a student team shooting, editing, and producing a video in the 24 hours of the Hackathon, recapping the event:

2016: https://www.youtube.com/watch?v=CMHpWSEIsVs
2015: https://www.youtube.com/watch?v=iRth69XrecY
2014: https://www.youtube.com/watch?v=vLFE5nKHSYI
2013--15 summary ("3 years in 3 minutes"): https://www.youtube.com/watch?v=akjFxR4sShc

More details about the OHI/O Informal Learning program are at:
http://arnab.org/about/ohio-ohio-states-hackathon-program

http://hack.osu.edu

Audrey Le, PhD Candidate Anthropology (Emphasis on Media and Education)
Teachers College Columbia University, al2812@tc.columbia.edu

I am interested in the **theoretical space of hackathons**, particularly in the **applications to industry contexts** and what related activities go by different names. Based on archival research and follow-up interviews with hackathon participants and organizers, I describe the ecological histories of hackathons in three industries and compare goals for hosting hackathons in terms of industry-specific preferences for the tools they produce for work.

Civic Hackathons

Hacking iCorruption was a forum for fellows at the Harvard Center for Ethics to present their research on institutional corruption, and get technical help to build tools to automate that research. Its sister event, Hack 4 Democracy, was organized by Team Democracy to connect technologists – those looking to volunteer their skills for a cause, with civic organizations in New England looking to hire that talent after the event. The goal of the first civic hackathon in Boston was to bring attention to the issue of money in politics. The legendary figures of law professor Lawrence Lessig – the founder of MayDay, the superPAC to end all superPACS, and then engineer student and activist Danny Miller drew six nonprofits fighting the same issue, and over 100 developers to both events.

The problem of technology in social movements was verification, or trust in prompting voters to take a stand on an issue. The TownHall team, for instance, wanted to do more than play with disapproval ratings, and looked to introduce the capacity to mandate new policies. The many hacks that appeared at Hack 4 Democracy were in fact incorporated into the workflow of the nonprofit sponsors, which in this case, were allowed to pitch and recruit help on existing projects.

Hackathons are a very recent affair. The first term was coined in 2006, but resembled code-fests where a critical mass of developers were needed to program a specific feature. Hackathons were only adapted to industry in 2012, starting with the Water Hackathon. Inspired by the steady rise in technical volunteerism in response to global humanitarian crises, the World Bank used the Random Acts of Kindness model to raise awareness of water-related challenges and create novel applications to improve infrastructure in resource-constrained areas. What civic hackathons aim to accomplish by hosting their events to crowdsource the expertise they are lacking to solve a problem of a relatively larger scale than other hackathons. In first countries, the hackathon is replaced by the yearlong challenge.

It should be noted that although the World Bank credits LinkedIn and Facebook for popularizing hackathons (2012, p. 2), hackathons of a more civic nature are organized around new developments in data science. For instance, the 2008 Apps for Democracy was sponsored by the city of Washington D.C. to mark the creation of a municipal open data catalog (Johnson and Robinson, 2014, p. 351). With the rise of open data, an alternative to the procurement process then existed for outsourcing government technology/software without direct oversight by the government except in the provision of data (Brahbam, 2009; Janssen et al, 2012).

Hackathons were also playing a growing role in the developing world. In July 2014, the Massachusetts General Hospital's Consortium for Affordable Medical Technologies (CAMTech) co-organized a hackathon in India called the Jugaadathon. 56 healthcare hackathons were held in 17 countries outside the United States between 2010 and 2014. Two organizations have helped aid

this expansion: Health 2.0 hosted code-a-thons in the Netherlands and China in 2012, and the Canadian organization Hacking Health held hackathons in Switzerland, Hungary, and Sweden in 2014.

Healthcare hackathons respond to a wide range of innovations in hardware and follow the startup weekend model. Participants come for the free business and technical mentorship and to win office space. An accelerator program provides winners with seed funding, a lean curriculum and a network of providers into which they can tap to validate market-fit of their hack. Healthcare hackathons aim to produce startups.

As a Chief Economics Commentator Grep IP reported in the Wall Street Journal this month, the healthcare industry, especially Big Pharma, had been lagging behind in large-scale innovation. In the age of fitness trackers, the founder of the MIT Hacking Medicine club, venture capitalist Zen Chu, claimed that there was no better time to disrupt the industry. The Associate Director at the Massachusetts General Hospital's Healthcare Transformational Lab, Maulik Majmudar, observed a significant drop in the use of fitness wearables like Fitbit after six to eight months. Research by Rockhealth indicated that most users had very low social and medical complexity; the problem, according to the director was not with the limited health functionality of wearables, but that they needed proper calibration to respond to the special needs of users with chronic diseases. At the 2016 Startup Week in Seattle, Brad Younggren MD proposed that artificial intelligence scripts, rather than telehealth solutions, could solve for simple triage questions to focus on quality care.

Journalism hackathons

Hacking Journalism NYC was organized by Hacks/Hackers, a data journalism education nonprofit started in 2009 under the auspices of the Knight Foundation, in partnership with the MIT Media Lab and Embed.ly. Embed.ly had been in the business of engaging with journalists by providing them with tools to interact with their reader base.

A recent New York Times innovation report had been leaked, highlighting the in-house need for collaboration between developers and journalists in order to compete with BuzzFeed and Huffington Post, whose model for digital journalism was far more effective at mastering new media tools. The Hacking Journalism 2014-2015 series were themed after new media: mobile, video and data science, its goals to identify the next trends in journalism and give an opportunity to data journalists to pick up new skills. Most importantly, the hackathons were used to support journalists in building tools that they would use to curate, not write, the news.

The hands-on professional development opportunity is most welcome in an era of bite-sized, user-curated news. Journalists were experimenting with models for editorial judgment, knowing that fake news was a big problem. Having lost advertising revenue to Facebook, news producers also looked to video and now augmented reality to offer an immersive experience that engage readers' affective response to the news. In mid-2014 when I started fieldwork, Hacking Journalism teams were just figuring out the script to automate the selection of the most important content in a written article. Opposition from more conservative reporters was futile; the National Association of Computer-Assisted Reporters (NICAR) was growing in membership. The new generation of tech-savvy journalists acknowledged that print and digital required different publication schedules and strategies.

Keywords: History of Hackathons, Digital Economy, Tools for Work

Power-Hour: A Case Study For Time-Bound Events On Crowdsourcing Platforms

Vassilis Javed Khan
Industrial Design Department,
Eindhoven University of
Technology
5612AZ, The Netherlands
v.j.khan@tue.nl

Bas van Hoeve
Design2Gather.com
Victoriapark 688, Eindhoven, The
Netherlands
bas@design2gather.com

Abstract
Time-bound events such as hackathons, installfests, designathons and other maker time-bound events are growing in numbers and diversity of topics. At the same time, online collaboration platforms are on the rise. In this position paper we report our preliminary ideas on how to design time-bound events in crowdsourcing platforms based on a case-study of "Power-Hour" an online, time-bound event that occurred twice in the month of October 2016 at Design2Gather.com.

Author Keywords
Crowdsourcing; Design; Time-bound Event; Online collaboration.

ACM Classification Keywords
K.4.3. Organizational Impacts: *Computer-supported collaborative work*.

Introduction – Particular Interest to the Workshop's themes
From the very first hackathon organized, in the end of the last century [1], the idea of hosting a time-bound event in which people actually make something at the end of it has taken off. This concept's popularity is evident by its diffusion into all continents but also into

other areas such as design in which designathons are hosted around the same concept.

Another growing development that has happened more recently is a new form of collaboration, in this case not on a specific physical space, but online, known as crowdsourcing. Although there are different definitions of what crowdsourcing is, we prefer to cite the following, encompassing definition: "*crowdsourcing is an umbrella term for a variety of approaches that harness the potential of large crowds of people by issuing open calls for contribution to particular tasks*" [2]. This definition includes paid crowdsourcing but does not restrict other platforms such as social networking systems or other computer supported cooperative systems.

Our specific interest for this workshop is the intersection of time-bound events and crowdsourcing. So specific to the listed themes, we are interested in the application of time-bound events in the context of crowdsourcing. We already know from research on online communities that events help the community to define itself by reminding members what they have in common and what their community is all about [3]. In the context of crowdsourcing platforms, Quirky [4] is a known example that has events at its core. Quirky organizes every week a time-bound event that members of its platform would join it –both online and on a actual space- and decide on a product idea that then the company would actually develop. Furthermore, OpenIDEO encourages its crowd to form "chapters" and then hold events on certain cities around the world [5].

Although the concept itself is not new, there is very little information in existing literature on how to design, organize, execute and evaluate such events on crowdsourcing platforms.

We would like to contribute to this aim by sharing our experience with "Power Hour", a time-bound event that was organized twice on the crowdsourcing platform: Design2Gather.com. Design2Gather.com is a crowdsourcing platform in which ideation and early design of specific tasks (known as "designments") is taking place.

On Wednesday the 5th of October from 11:30AM - 12.30PM (GMT) and Thursday the 18th of October from 12:30AM - 13.30PM (Figure 1) Design2Gather.com organized two "Power-Hour" events on its platform. Power-Hour was aimed at boosting the crowds' designment process. During this event, Design2Gather employees were available to proved crowdworkers direct feedback on their ideas, inspiration and deliverables for the designment: "Redefine Kitchen Lifestyle". In that event, workers were able to brainstorm about the project with other designers and ask more detailed questions about the project.

The designment's description was:

How can you make cooking or anything around the kitchen easier, smoother or just more fun? An international brand that competes with the big kitchen brands in the world wants new innovative kitchen items! We are looking for products that have stainless steel as the main material but this can be combined with other materials to create a clever new product that makes your life easier in the kitchen. You are not limited to redesign existing items such as spoons, forks, knives, cups, mugs, bowls, pots and pans. Original new ideas with combined functions are

encouraged! Durability and aesthetics are important as well and the product should express a sense of quality and multi-functionality. Are you hungry to create fresh, new and innovative kitchen tool ideas? Join now!

During the workshop we would like to share the data we logged regarding the involvement of workers as well as our experiences in how to organize, design, execute and evaluate this type of time-bound events.

Motivation & Background

Our motivation is twofold: 1) we would like to share our experience in organizing a time-bound event in a crowdsourcing platform and 2) we would like learn from other workshop participants good practices of time and space-bound events that could transfer to the context of crowdsourcing.

Dr. Vassilis Javed Khan is assistant professor at the Industrial Design Department of Eindhoven University of Technology in the Netherlands. His currently research interest and activities can be captured in the following title: *Design for Crowdsourcing & Crowdsourcing for Design*. What is mean with the first part is how designers can improve existing crowdsourcing platforms and by the second part what is meant is develop novel crowdsourcing platforms to support designers.

Bas van Hoeve is an alumni of the Industrial Design Department of Eindhoven University of Technology in the Netherlands and the creative director of Design2Gather.

Figure 1: Online ad that promoted the "Power Hour" event on Design2Gather.com

References

1. Hackathon. (n.d.). In *Wikipedia*. Retrieved December 15, 2016, from https://en.wikipedia.org/wiki/Hackathon
2. Geiger, D., Rosemann, M., Fielt, E. and Schader, M. Crowdsourcing information systems-definition, typology, and design. Thirty third International Conference on information systems, (2012).
3. Kim, Jo, A. (2000). *Community building on the Web*. Peachpit Press.
4. https://www.quirky.com/
5. https://beta.openideo.com/chapters

Chris Holdgraf
Data Science Fellow | Graduate Researcher
Berkeley Institute for Data Science, UC Berkeley
choldgraf@berkeley.edu | github.com/choldgraf

The landscape of tools for performing research has changed considerably over the past few decades. The increased use of technology, the ease with which we can share ideas with one another, and the growth of data-centric methods have all influenced how we ask and answer questions about the world. However, as these tools have grown quickly, our ability to understand best practices to use them has been slower to evolve. The fundamental approach to performing research and teaching in universities has remained relatively unchanged over the past century – graduate students and post-docs spend much of their time working independently (one might say working in isolation), and they rely on a small group of colleagues such as a lab for guidance and community during this time.

While academia has stuck with this model for quite some time, there are many other ways to organize people in order to accomplish a goal. My interest in this hackathon stems from the hope that academia can explore alternative models for working in teams to get things done. I've seen this work well in previous projects that I have been involved with. For example, during graduate school I worked with a team of graduate students to create the *Beyond Academia* organization and conference [0], which teaches graduate students about their options outside of the traditional academic path. I was also a team leader on a project to create teaching materials for graduate students learning to analyze data more effectively. This is now being turned into a more general *Open Course in Neural Data Science* [1].

As I finish graduate school, I've recently begun a fellowship at the Berkeley Institute for Data Science (BIDS), an organization that sponsors efforts to explore new models for scientific inquiry and education. In line with the goals of this workshop, I am currently organizing a hackathon-style event to be held in early March. The event will center around writing documentation, tutorials, and guides for open-source software, as well as building tools that enable people to make their software more intuitive for users. We're calling it a *Docathon* [2], and are organizing simultaneous working groups spread out across the world.

While there is much enthusiasm around hosting this event (and others like it), the organizing team still has much to learn about running hackathon-style sessions as effectively as possible. This "meta-hackathon" workshop is a great opportunity both to share my experience with others, and to learn best-practices for our specific event. I hope that it will also serve as inspiration for organizing future events at UC Berkeley. Below are a few specific ideas that I'm interested in discussing at the workshop:

I would love to see discussion about how hackathons can exist as a part of a larger arc of collaboration between groups of people. For example, many projects begin with a hackathon-style event, and this serves to build momentum and energy that fuels work that is then carried out over the next several weeks. This seems particularly useful for a university setting in which

projects have a natural life-cycle of one semester, but where team members may come from disparate backgrounds and fields. An event that brings everybody together, crystallizes goals, and clarifies how team members complement one another could be extremely useful in building new models for team-based projects in research and education.

I would also like to explore the extent to which hackathons may be useful as a beacon to attract new contributors and team members, and to more generally reach out to individuals that may be interested but need direction. One challenge that organizations like BIDS face is that they serve a niche in the university that many researchers don't realize exists. It is a challenge to embed the kinds of services and work that BIDS offers into the broader university culture, and events are often a good way to bring people together. I'd like to meet other hackathon organizers to see whether they've been able to grow their organizations and networks by holding hackathon-style events.

Finally, an overarching question that I'm concerned about is how to structure these events such that they are sustainable and repeatable, and don't demand an extreme amount of commitment from a single person. Universities tend to be composed of teams that are in constant flux, and it is challenging to create group structures that are able to withstand the test of time. One reason for this is that organizations and events are often spearheaded by a single, highly-interested, strong-willed person that makes things happen. However, once that person leaves there is not enough willpower to fill the vacuum that is left behind. I'd love to see discussion about how to spend time thinking about the high-level structure for events like hackathons (e.g. how to spread responsibility, document what works / doesn't work, and transition between organizing teams) such that they are able to exist beyond one or two iterations.

I think that all of these topics are directly related to the goals of this workshop, and I'd love to participate, share my own knowledge, and hear from others over the course of the day. My hope is that hackathons, and short-term projects more generally, will become a larger part of the toolkit universities have for getting things done. I think that participating in the hackathon workshop will give me valuable experience in order to make this happen.

[0] beyondacademia.org/
[1] github.com/neuro-data-science/neuro_data_science
[2] github.com/BIDS/docathon

Time-bounded Collaborative Events: An Organizing Framework and Applications

Anna Filippova
Institute for Software Research
Carnegie Mellon University
Pittsburgh, PA 15213, USA
annafil@cmu.edu

Erik Trainer
James D. Herbsleb
Institute for Software Research
Carnegie Mellon University
Pittsburgh, PA 15213, USA
{jdh,etrainer}@cs.cmu.edu

CSCW '17, February 25 - March 01, 2017, Portland, OR, USA
ACM 978-1-4503-4688-7/17/02.
http://dx.doi.org/10.1145/XXXX

Author Keywords
Time-bounded collaborative events; hackathons; sprints; codefests; edit-a-thons; open-source software; scientific software; radical collocation; innovation; collaborative learning.

ACM Classification Keywords
K.4.3 [Computers and society]: Organizational Impacts–Computer-supported cooperative work; H.5.3 [Information interfaces and presentation (e.g., HCI)]: Group and Organization Interfaces–Computer-supported cooperative work

Introduction
Time-bounded collaborative events where teams work under intense time pressure have become popular interventions to stimulate problem solving, informal and collaborative learning, and community growth. The problems targeted, the participants selected, and the objectives of organizers can vary widely. The format of the event, technology support, and team configurations and work processes can also vary widely. This means that one event may differ from the next along many dimensions; it is unlikely that organizers can wholly borrow from previously run events given the unique needs, resources, and constraints of their communities. It would therefore be very helpful to have a principled approach to the various decision-making points involved in planning and executing a successful hackathon.

Figure 1: Organizing framework for time-bounded collaborative events.

Drawing on our own experience studying hackathon-like events and informal conversations with organizers, we propose an organizing framework (Figure 1).

In this model, the community operates within a particular social context and around particular technologies. It supplies inputs, such as participants and funding, to an event designed to enact change. The event produces outputs, which are absorbed by the community over time through impact. Outputs are immediate and small in scale compared to impact. For example, an output of an event may be at the scale of individual participants learning new tools and techniques. The eventual impact of that event may be a community with a stronger collective technical capability, expressed through subsequent completed projects and collaborations.

- **Goal.** What problems or needs is the community trying to address? This would include such things as accelerating project development, learning new tools and techniques, building social ties, or enhancing awareness of what people are working on. The goal will influence the **inputs**, or resources the community supplies. For instance, learning new tools suggests a need for a pool of newcomers and mentors to help them.
- **Context/environment.** What are the characteristics of the community and the technology around which it operates? This would include aspects of the *technology structure* (e.g., architecture, programming languages, dependencies), *community structure* (e.g., range of application domains, skills, roles), *community climate* (e.g., maturity, needs and issues), and *domain area* (open-source, science, collegiate, company-internal). Context influences the **goal**. For instance, software maintainers finding it difficult to meet users' needs may benefit from an event that teaches users about the codebase and how to contribute. The context will also influence the **inputs**. For example, there needs to be sufficient interest among appropriately skilled users to hold such an event.
- **Inputs.** What raw materials will be drawn on to execute the event? Materials might include the *participant pool* (geographic location, areas of expertise, incentive structures), an appropriate *time span* for the event, *sponsorship* needed to pay for the space, prizes, and food, and *institutional constraints* such as funds available to participants for travel costs, and the

types and brands of products vendors can provide. Inputs are a product of the **context/environment**, and constrain the available choices when designing event **activities**.

- **Activities.** What will be done with the inputs to direct the course of the event? Activities can be divided into **baseline activities** essential to the event, and **variable activities** that have multiple known values (shown in Figure 1 with an asterisk). Baseline activities include such things as *creating an agenda* (e.g., training workshops all morning and hacking in the afternoon), *providing technology support* (e.g., version control, testing and integration tools, communication channels, logging progress), having a designated hackmaster to *facilitate* hacking (e.g., making introductions, enforcing rules), holding *social events* that afford informal interactions, and *evaluation* (e.g., finding judges who have expertise and represent the teams' values, selecting appropriate prizes). Variable activities include such things as *participant selection* (first come first served, selection by computer algorithm, semi-curated, fully curated), *preparation activities* (making introductions, pitching ideas, reading up on related work), and *configuring team structure* (new/existing, collaborative/competitive, task interdependence, face-to-face/remote, homogeneous/heterogeneous). Activities influence, and are influenced by, team processes. Selecting participants from distinct subgroups will likely require bridging work before they can work together since they do not have a common vocabulary or shared understanding of the work. There may be more conflicts and misunderstandings compared with groups where there are no such perceived rifts. On the other hand, highlighting overarching goals and establishing constructive and inclusive brainstorming flows can positively influence activities [2].
- **Team processes.** What are the work practices and social dynamics that influence activities and outputs? Team processes include forming, handling *conflict*, developing group *norms*, and *brainstorming*. If a team agrees to use GitHub to conduct their work, for instance, it is important for all team members contributing code to use it. Otherwise, the team will likely need to expend extra effort to integrate everyone's work, which distracts them from finishing their work in time.
- **Output.** What outputs arise from event activities and team processes? Output might include such things as having a *fun* and enjoyable experience, the creation of a *prototype* or proof-of-concept, *new contacts* made, *learning* new tools or skills, being able to make *contributions* to a new codebase, *awareness* of who knows what or who has what skills in the community, or seeing participation from a *diverse audience* (along dimensions such as gender, self-efficacy, age, career stage, ethnicity).
- **Impacts.** What are the changes in the community, in the short or long-term, that result from event specific outputs? Examples of short-term changes might include *research products* like publications, working groups, and proposals, and *software products* like new tools and datasets. Longer-term effects might include *expanded collaboration networks*, increased *community diversity*, increased *technical capabilities of community members*, and higher *contribution rates*. Impacts are likely to feed back to inputs. For example, an event where people build up knowledge of who knows what, and expand their networks, would result in a more visible participant pool that could be leveraged to advance work on specific development tasks.

Using the model

In our presentation, we will describe examples of detailed, step-by-step uses of this model to impact various types of goals for event organizers. Beginning with a set of event goals and constraints, we will illustrate how to select and evaluate relevant event outcomes and long-term impacts for organizers to track, as well as how to identify important inputs that may shape event design. We will also explore examples of supportive baseline and variable activities that can be designed to support event goals within these constraints. Finally, we will consider possible interactions with intra- and inter-group processes that may either support or conflict with event goals.

Example: Bridging communities

Below we present an example flow through the model.

Event goal: To establish collaborations between several distinct communities that presently are largely unrelated through social networks or shared tools.

Context: A scientific software community with a code base that has potential to impact a variety of scientific fields and an average learning curve.

Relevant event outcomes:
Establishing new contacts, enhancing awareness of other communities' needs, event attendee diversity.

Sample evaluation metrics (such as by using a survey):

New contacts: "Of the contacts you made at this event, how many will you pursue NEW projects with in the future? (Projects could be developing new open source tools, research collaborations, a new jobs or others)" [Numeric entry]

Awareness[#]: After this event, I feel that I have a better understanding of: [Strongly Disagree-Strongly Agree]

- Which [team / community*] members have expertise in specific areas
- What tools/methods are predominantly used within other [teams / communities*]
- What other [teams / communities*] are currently working on
- The issues and challenges faced by other [teams / communities*]
- Common workflows in other [teams / communities*]
- What open research questions other [teams / communities*] have (if applicable)

* [replace group noun as applicable] # Measurement scale currently under development

Attendee diversity: Demographic questions that include current education level, field of employment/role, years of programming experience (or a programming self-efficacy scale).

Please refer to this link for a sample post-event survey (as evaluated in [2]).

Possible relevant impacts:
Cross-disciplinary research outcomes, increased project community diversity, higher contribution rates.

Sample evaluation metrics (at suitable time intervals post event, such as one month, 6 months, 1 year):

- *Research outcomes:* citation counts and topic areas of work citing the software package (6 months to 2 years, depending on field)

- *Project community diversity:* number of new issues created, number and type of new features requested or added, mailing list/other virtual channel discussion volume and subject area (1 month to 1 year)
- *Higher contribution rates:* pull requests submitted and accepted over time (1 month to 1 year)

Inputs:
Given the event goal, the target pool of participants consists of multiple distinct communities that may be distributed across multiple geographic locations, with potentially different motivations and incentive structures. Therefore the amount of time available to bring these communities together would likely be limited - several days or a week, depending on the amount of funding support available. Additional questions to consider may be: are the tools necessary to work with the software project easily replicable across sites, or is particular and specialized hardware required? What types of data sources does each community work with, and what additional work is required in order to make these sources cross-compatible?

Mediators/Processes:
Time-bounded intensive collaborative events that feature distinct communities may face several intra- and inter-group challenges. Social psychology research suggests that constructing teams that are made up up distinct communities (such as teams of participants from different fields or research backgrounds), may create tension and introduce communication challenges [7]. This is because working together on a shared task may highlight hitherto invisible differences in background or experience. These differences may create a perception of a distinct “in-group” of like-minded participants, and an “out-group” of participants that are unlike oneself. These differences can further be misinterpreted as lack of competence or interest in the ideas of the other group, resulting in interpersonal friction and potentially outward manifestations of conflict [5].

This may be particularly true if a team is composed of two distinct homogeneous groups, compared with an entirely heterogeneous team in which every member has a different background from other team-members [1]. Multiple visible differences that line up along the sub-group boundary, known as faultlines, may exacerbate these effects. Examples of faultlines may include participants from the same fields of inquiry also hailing from the same geographical area or institution, thus having potential pre-existing relationships and ability to communicate with each other prior to the event [4]. Another example may be sub-groups representing different fields that also have homogeneous levels of experience within each sub-group but different between groups, such as a sub-group of early career researchers from one field, and a second sub-group of later career researchers from another field.

Relevant Activities:
Given the above constraints, below is a list of a possible set of important activities to focus on in order to address event goals and counteract the above challenges.

Selecting participants that best represent their respective communities: In this situation, a semi-curated or wholly curated participant selection strategy may be the most appropriate to ensure representation of important communities/groups at the event, as compared to a first-come-first-serve sign-up method. A curated participant selection strategy would involve inviting specific members of a community only, while a semi-curated method may include first-come-first-serve sign-ups alongside invitations. To encourage awareness of different community needs, communities of interest should be identified and participants should be selected who are likely to be aware of and represent major

community issues. Research suggests selecting community members who are central to the social network of their respective communities (for example, researchers whose work is highly cited and relevant to the themes of the event) [3]. It may be relevant for the event goals to also select participants with different levels of experience (either at different stages of their research careers or software development experience). A precise call for participation and set of selection criteria can facilitate this.

There are also distinct activities that may be used to counteract possible challenges of sub-group dynamics, given the participant selection above. Emphasizing interdependence of goals, and shared outcomes across different participant groups may support the establishment of a common community identity in lieu of subgroup identities [6]. This can be achieved, for example, through an introductory session at the start of the event that identifies common benefits of collaboration across different sub-groups. Mixers at the start of the event that encourage individuals to sit next to and interact with participants they have not met before can serve to temper the effects of pre-existing ties among participants. Such mixers can be targeted at developing an initial set of goals or an agenda for the coming day(s) of the event and support formation of teams with individuals from different backgrounds. After teams have been formed, the use of brainstorming strategies such as focusing on integrating ideas generated, avoiding criticism, and allowing all ideas to be discussed, no matter how freewheeling have been shown to improve well-being of diverse teams [2]. Finally, a variety of social mixers before, during and after the event (such as common lunches and dinners, shared modes of transportation), can support establishing stronger ties among participants from different communities.

The above strategies presented address the specific concerns that an event with this example goal may face, however there are a number of other baseline event activities that will likely need to be performed by organizers such as securing venue, ensuring adequate technology support, and so on. These are not discussed here due to space constraints.

What we hope to achieve from the workshop

We aim to draw on conversations and presentations from the workshop day to refine the above model, and understand to what extent our predictions are relevant, and what elements may yet be missing and need to be investigated.

Author backgrounds

Anna Filippova (*primary contact*) is a postdoctoral researcher with the Institute for Software Research at Carnegie Mellon University, where she studies the role of events in supporting open collaborative community development. She has several years of experience in organizing open-source community events, including large-scale conferences like Abstractions and Red Dot Ruby, and monthly meet-ups. Her PhD work with the National University of Singapore examined the impact of different forms of conflict on Free and Open Source Software development. She has also studied group norm evolution and normative conflict in virtual spaces and open collaborative communities.

Erik Trainer is a post-doctoral researcher in the Institute for Software Research at Carnegie Mellon University. He received his PhD in Information Computer Science from the University of California, Irvine in 2012. His research focuses on creating technologies and practices that support the relationships of people engaged in technical work, especially in open-source software development and software production in science.

James Herbsleb is a Professor of Computer Science at Carnegie Mellon University, where he serves as Director of the PhD program in Societal Computing. His research interests focus on global software development, open source, and more generally on collaboration and coordination in software projects. He was recently awarded the SIGSOFT Outstanding Research Award in 2016, and previously the Alan Newell Award for Research Excellence in 2014. He has served on the PC of several conferences, including ICSE and FSE, was co-chair of CSCW 2004, and served as an associate editor of ACM Transactions on Software Engineering and Methodology.

REFERENCES

1. Andrew M Carton and Jonathon N Cummings. 2012. A theory of subgroups in work teams. *Academy of Management Review* 37, 3 (2012), 441–470.
2. Anna Filippova, Erik Trainer, and James D. Herbsleb. In Press. From Diversity by Numbers to Diversity as Process: Supporting Inclusiveness in Software Teams with Brainstorming. In *Proceedings of the 39th International Conference on Software Engineering (ICSE 2017)*.
3. Julija Mell, Daan van Knippenberg, Wendy P van Ginkel, and Pursey Heugens. 2014. Whose Brain to Pick? Boundary Spanning and Transactive Memory in Inter-Group Knowledge Integration. In *Academy of Management Proceedings*, Vol. 2014. Academy of Management, 12519.
4. Jeffrey T Polzer, C Brad Crisp, Sirkka L Jarvenpaa, and Jerry W Kim. 2006. Extending the faultline model to geographically dispersed teams: How colocated subgroups can impair group functioning. *Academy of Management Journal* 49, 4 (2006), 679–692.
5. Tony L Simons and Randall S Peterson. 2000. Task conflict and relationship conflict in top management teams: the pivotal role of intragroup trust. *Journal of applied psychology* 85, 1 (2000), 102.
6. Henri Tajfel. 1982. Social psychology of intergroup relations. *Annual review of psychology* 33, 1 (1982), 1–39.
7. Henri Tajfel and John C Turner. 1979. An integrative theory of intergroup conflict. *The social psychology of intergroup relations* 33, 47 (1979), 74.

The CHI4Good Day of Service: What is Produced?

Emily Porter
University of Colorado
Boulder, U.S.A.
Empo0270@colorado.edu

Elizabeth Gerber
Northwestern University
Evanston, U.S.A.
Egerber@northwestern.edu

Christopher Bopp
University of Colorado
Boulder, U.S.A.
Chris.bopp@colorado.edu

Amy Voida
University of Colorado
Boulder, U.S.A.
Amy.Voida@colorado.edu

.

Abstract
The popularity of hackathons has increased as technology pervades more diverse facets of our lives. Traditionally designed for computer programmers, hackathons are now being appropriated by new stakeholders. Yet with this evolution in hackathons, we no longer adequately understand what is produced and, thereby, the value of these events. We conducted an interview study with 22 stakeholders—participants, representatives of nonprofit organizations, and organizers—of the CHI4Good Day of Service to understand what is produced through philanthropic hackathons.

Introduction
Historically hackathons have been recognized as a site for the development of prototypes in the form of code or other physical artifacts. More recently, however, diverse stakeholders across sectors appropriate hackathons to address their own challenges, resulting in hackathons newly-dubbed as "civic" [3] or "social-issue" hackathons [6]. These new styles of hackathons have been described as "addressing social conditions and their consequences" [3]. With this evolution in hackathons, however, it is no longer clear what the hackathon genre produces.

These philanthropic hackathons are generally understood by researchers as more complex forms of the conventional hackathon. Multiple researchers have independently observed that philanthropic hackathons accomplish more than the physical outcomes typically

About The Authors

Emily Porter: Emily is a MSc candidate with the ATLAS Institute at the University of Colorado, Boulder. Her work focuses on technology for developing communities.

Christopher Bopp: Chris is a Ph.D candidate with the ATLAS Institute at the University of Colorado, Boulder. He conducts research in human-computer interaction and computer supported cooperative work, with a focus on philanthropic informatics

Elizabeth Gerber: is an Associate Professor of Design at Northwestern University. She is interested in broadening and diversifying participation in innovation work to increase the range of solutions tackled. She has organized many civic hackathons and would like to deepen her understanding of the practice.

Amy Voida: Amy is an assistant professor in the Department of Information Science at the University of Colorado Boulder. She conducts research in human-computer interaction, computer supported cooperative work and ubiquitous computing, with a focus on philanthropic informatics.

associated with traditional hackathons, suggesting that these events also provide a venue for knowledge exchange, public relations opportunities, and citizen that philanthropic hackathons engender new forms of production work prompts the question: what is the breadth of production work when traditional hackathons are appropriated in a philanthropic context?

Our research seeks to understand how production work changes when the traditional genre of a hackathon is appropriated in philanthropic contexts. Through interviews with 22 stakeholders of the CHI4Good Day of Service, we learn that in addition to artifacts, philanthropic hackathons produce technical expertise, expanded social networks, an exposure to design process, affective experiences, and an opportunity for participants to shape their identity against a cross-sectoral, interdisciplinary backdrop

Results: The Production Work of the CHI4Good Day of Service

Artifacts

True to the traditional hackathon, most participants noted contributing to a digital artifact that was handed over to the nonprofits after the Day of Service concluded. While a few of these artifacts were prototypes (e.g., a digitized searchable "re-entry" guide for those exiting the American prison system) most common to a traditional hackathon, they were the minority of digital artifacts created during the Day of Service. More frequently teams implemented small improvements for the nonprofits' existing websites or they produced wireframes and visual mockups. In some cases the artifact was actually a draft document detailing important "next steps" for the NPOs. Yet, despite the broad-based interest in producing digital artifacts, participants reported mismatches in expectations and needs, particularly between the volunteers and nonprofit representatives.

Technical Expertise

The lack of technical expertise among nonprofits representatives going into the event was a central theme among both volunteers and nonprofit representatives. Almost all nonprofits interviewed discussed being unsure of what skills or tools would be needed to accomplish the project they had proposed for the event. Most agreed that participation in the event led to a better understanding of their baseline project needs, awareness of existing technologies, their potential capacity to serve these needs, as well as a basic level of competency in determining what skills might be needed for their projects in the future. In a minority of cases, participants suggested that technical expertise was not produced for the nonprofits in a way that would cultivate any kind of long term impact for the organization.

Design Process Experience

Most projects proposed for the event were ill-suited for either the timeline of the day or the expertise of the volunteers. Most projects, then, required restructuring or adaptation; in some cases the entire proposal was scrapped by volunteers. As teams restructured project proposals or, in some cases, developed totally new ones, there was a discernable shift in methods reported by nonprofits to reshape the plans. In one case the user-centered design processes used during the event were described as an impetus for change within the organization.

Social Networks
Individuals in all stakeholder groups were motivated to attend the event by the potential to enhance their social network. Ties were developed between nonprofit organizations and volunteers, among volunteers, between organizers and nonprofit organizations, and between organizers and volunteers. Participants found different kinds of ties more or less valuable based on perceived need at a personal or organizational level. For example, same valued the opportunity to meet other volunteers. Participants believed that these new network ties would lead to possible job or volunteer opportunities, new volunteers, as well as the potential for extended intercommunication between disciplines and backgrounds.

Affect
Affect emerged as a more abstract product of the Day of Service—an unexpected but significant theme across stakeholder groups. Both volunteers and nonprofit representatives alluded to emotional changes throughout the day brought about by social interactions, event logistics, and perceptions of progress (or the lack thereof). Terms like "energy," "fun," and "good feeling" were used to describe positive affect while "awkward," "uncomfortable," and "frustration" were used to describe negative affect. Many individuals mentioned that their intention to participate in the event was largely motivated by the idea that some sort of positive affect would be produced.

Hackathon Identity
The work of navigating the diverse backgrounds and varied levels expertise of team members within each hackathon group afforded an opportunity for individuals to explore and construct their hackathon identity, much in the same way that Arrow and McGrath find that processes created in small group settings are fundamental to establishing the identity of the group as well as the individuals within the group [2]. By understanding the goals, motivations, and other information about team members, participants were able to work out their identity with respect to the role they might play and what they might contribute to the project at hand.

Numerous participants noted that skill matching with projects was of particular importance to the production of their hackathon identity. Additionally, unlike typical hackathons, volunteers at the Day of Service were able to rotate through the event based on personal preference or schedule. Participants commented that the fluid group composition created a context for the continual renegotiation of one's hackathon identity.

Discussion
The goal of this research was to understand the value of philanthropic hackathons to the CHI community and their community partners. Results of this research suggest that these events not only produce digital artifacts, but also technical expertise, expanded social networks, an exposure to design process, affect, and occasions for shaping cross-sectoral, interdisciplinary identities.

Irani suggests that hackathons may favor "quick and forceful action" over "the slow construction of coalition across difference" [5]. Gregg raises similar concerns about whether hackathon-style events may move too quickly from complex social issues to overly simplistic technical solutions. Given these important concerns, we

turn to consider how we might better design future philanthropic hackathons to more deliberately help form and foster meaningful connections among the stakeholders of the event and to encourage teams to engage more deeply with social issues, moving less quickly, if at all, to technical solutions.

Design Implications for Philanthropic Hackathons

Based on our empirical data, we speculate that the hackathon genre might be productively restructured to benefit the breadth of stakeholders and suggest a number of implications for the redesign of philanthropic hackathons. Organizers might better support technical capacity building and expertise by offering design patterns, or general templates—e.g., for setting up databases, creating websites, developing mobile applications—for solutions in a format that doesn't require specifics tied to a particular partner organization. Organizers might better support the expansion of social networks by connecting participants before the hackathon to communicate their intention to collaborate and after the hackathon as a record of their collaboration. To support an exposure to design process, organizers might offer step-by-step guidance through a flexible design process, producing user needs' assessments and workflows prior to attendance at the event. As affect has been found to accompany creative activity [1], organizers might support the expression of both positive and negative affect through opportunities for active reflection throughout the day. Lastly, to support occasions for shaping identities for collaboration, organizers might propose distinct roles and responsibilities for each participant to achieve shared goals.

Conclusion

Through this research, we have developed a richer understanding of how philanthropic hackathons go beyond creating digital artifacts—to support technical expertise, design process, social networks, affect, and identity. Understanding the value in the breadth of what is produced by philanthropic hackathons will enable us to rethink the design of these events and will, we hope, spark a conversation about how the HCI community can use best use hackathons for societal good.

References

1. Teresa M. Amabile, Sigal G. Barsade, Jennifer S. Mueller, and Barry M. Staw. 2005. Affect and Creativity at Work. *Administrative Science Quarterly* 50, 3: 367-403.
2. Arrow, H. and McGrath, J.E.. 1993. Membership Matters How Member Change and Continuity Affect Small Group Structure. *Small Group Research* 24, 3: 334-361.
3. DiSalvo, C., Gregg, M., and Lodato, T. 2014. Building Belonging. *Interactions* 21,4 : 58-61.
4. Gregg, M. 2015. Hack for good: Speculative labour, app development and the burden of austerity. *The Fibreculture Journal* 25, 186: 183-201.
5. Irani, L. 2015. Hackathons and The making of Entrepreneurial Citizenship. *Science Technology & Human Values* 40, 5: 799-824.
6. Lodato, T.J. and DiSalvo, C. 2016. Issue-oriented hackathons as material participation. *New Media & Society* 18, 4: 539-557.

Title	Organizing community coding events in open source biological research
Author	Brad Chapman
Affiliations	Harvard Chan School Bioinformatics Core (http://bioinformatics.sph.harvard.edu/)
Contact	bchapman@hsph.harvard.edu
Keywords	bioinformatics, community, diversity, training

The Open Bioinformatics Foundation (https://www.open-bio.org) is a community of scientists creating open source code to solve biological problems. A yearly conference, started in 2000, provides the opportunity for in person discussion and presentation on technical work about code development and biological analyses.

In 2010, we recognized a need for a more practical hands on working session in addition to the conference and developed a two day coding session called the OpenBio Codefest (https://www.open-bio.org/wiki/Codefest). This event continued the past 7 years in a wide diversity of locations, with the most recent taking place at an Orlando makerspace (https://www.open-bio.org/wiki/Codefest_2016).

At the CSCW 2017 Hackathon workshop, I'll discuss the organization and evolution of Codefest. It initially started as a space for community members who were already collaborating remotely to sit together and work. Over time, it grew to better incorporate new members into the community by serving as a fun and open environment for sharing work and meeting like-minded researchers.

The positive things we've learned in organizing Codefest are:

- The value of collaboration over competition. Codefest has no prizes or competitive structure, and instead focuses on producing useful practical code that we can share at the associated conference and more widely through blog posts and scientific papers.
- The power of self-organizing groups. We do not pre-define the agenda for Codefest and let the attendees suggest areas of focus and then provide introductions so working groups can form. This allows newer community members to work alongside more experienced developers in areas they'd like to learn, and to allow the community to shift focus with new technologies and approaches.
- The advantage of in person discussion for developing interoperability standards. One successful outcome of Codefest have been the development of tool communication standards which allow different communities to share development resources. Like other projects at Codefest, standards creation happened organically due to the need for larger projects to be able to better to re-use analyses.

We also currently face challenges that we're hoping to learn potential solutions to:

- How to attract a more diverse set of community members. Like many programming and bioinformatics conferences, we struggle to attract a diverse crowd of attendees. As a result, Codefest can feel intimidating or unwelcoming to those outside the community. We've received universal praise that we're welcoming once overcoming that initial hurdle, but would like ways to project this welcoming attitude so under-represented researchers feel comfortable investing their time and expertise at Codefest.

- Incorporating teaching and training into the content of Codefest. As we've increasingly tried to attract new community members, we've developed the need to help integrate them into the community. In many cases, new members will be experts in some areas but not in the projects or languages under active development at Codefest. We need to develop methods to quickly get them comfortable and productive so they can contribute within a reasonably short time frame.
- Scaling events to incorporate new members and approaches. As we actively recruit new attendees we're running into the issue of figuring out how to support them at larger scale. Our approach of having a few mentors who make connections and provide orientation on projects will need improvement if we're successful in recruiting new, diverse attendees.

Attending the Hackathon Workshop is a chance to show areas where we've been successful, and to learn how to be better organizers. We hope to continue to expand and improve Codefest and related events for the open bioinformatics community.

CSCW 2017 workshop application: *Hacking and Making at Time-Bounded Events: Current Trends and Next Steps in Research and Event Design*

Jonathan Morgan, Wikimedia Foundation, jmorgan@wikimedia.org
Dharma Dailey, University of Washington, ddailey@uw.edu
Benjamin Mako Hill, University of Washington, makohill@uw.edu

Keywords

Collaborative learning; data science; hackathon

Abstract

Learning to program can be a valuable skill for people who only need to program occasionally, in specific contexts, and even for people who never engage in programming at all. The rise of data science as a profession and the proliferation of data by and about internet users creates a similar use cases for pedagogical approaches that support “end user” and “conversational” level literacy in that domain. We developed the Community Data Science Workshop, a four day, hackathon-style programming series, with the goal of democratising data science through a collaborative, hands-on approach focused on using open data resources to answer real-world questions. We discuss design decisions we made to support self-directed learning and exploration among people with no previous experience with programming or data science, as well as challenges we faced and questions raised by our experience leading these events.

Community Data Science Workshops

Nearly every published discussion of data science education begins with a reflection on an acute shortage in labor markets of professional data scientists with the skills necessary to extract business value from, burgeoning datasets created by online communities like Facebook, Twitter and LinkedIn. This model of data science – professional data scientists mining online communities for the benefit of their employers – is only one possible vision for the future of the field. What if everybody learned the basic tools of data science? What if the users of online communities – instead of being ignored completely or relegated to passive roles of data producers to be shaped and nudged – collected and analyzed data about themselves? What if, instead, they used data to understand themselves and communicate with each other? What if data science was treated not as a highly specialized set of skills but as a basic literacy in an increasingly data-driven world?

Most data science courses and programs focus on providing credentials and a full set of tools and techniques expected of professional data scientists. As such, they often require significant

programming expertise as a prerequisite. Although recent years have provided more and more opportunities to learn programming outside of computer science and engineering degree paths at colleges and universities, there are still significant barriers to learning data science, especially among people without time or money to pursue a degree in a related field. Additionally, most programming courses available, whether classroom or internet-based, focus on programming fundamentals and/or are geared specifically towards software or web development. While it is possible to master the skills necessary to programmatically collect, manipulate, and visualize large datasets without a background in computer science or software engineering, people interested in taking this trajectory are presented with few options.

Increasing data science literacy among the broader population calls for a different approach to programming education, in which programming is introduced as a means rather than an end in itself. It also requires a supportive environment that accommodates different skill levels and learning styles, and a curriculum that provides incremental value, encourages collaboration and experimentation, and ultimately gives students the just the right amount of insight into what's happening "under the hood" of the tools they use that they feel confident applying their new knowledge to questions that occur beyond the learning context.

Time-bound, collaborative events such as hackathons and editathon can be a useful model for teaching data science literacy, because they often fulfill similar purposes and operate under similar constraints. Between 2014 and 2016, we created and implemented one such event, the Community Data Science Workshop (CDSW) at the University of Washington. CDSW is a free, Python-based introductory programming series focused on increasing data science literacy among people with no previous programming experience. Like many other collaborative events, such as the Boston Python Workshop and the themed edit-a-thons run within the Wikimedia Movement, CDSW had a strong outreach component: our goal was to recruit and engage people who are less likely to have previously had opportunities to learn data science skills—women, students in the humanities and social sciences, and working professionals in non-developer/non-analyst roles whose jobs nonetheless involve working with data.

Design of the workshop

We ran the CDSW series, which consists of one Friday evening and three Saturday sessions, four times between 2014 and 2016. Like many hackathons, CDSW is open to anyone, provided at no cost to participants, and held on weekends in order to make it easier for people with full-time jobs to attend.

The workshops are heavily project-driven—students are shown how to retrieve and manipulate data through open APIs from online communities such as Wikipedia and Twitter and encouraged to formulate and answer their own questions. Throughout the workshop, mentors circulate freely across sessions and working groups, available to provide support and answer questions. We aim to provide a 4:1 student-to-mentor ratio in order to make sure a mentor is available any time a student got stuck.

The Friday session is devoted to helping participants install the necessary software and configure their personal computers, and become familiar with the command line environment. Each Saturday session begins with a two-hour interactive lecture in the morning that builds upon the topics presented in previous sessions. Each afternoon session is organized around open-ended questions designed to foster structured exploration of the morning's concepts to help participants synthesize and use their new skills. Afternoon sessions involve independent project work. Participants are given an archive of several simple programs written using only concepts that were introduced to in the lectures. After a short exposition and explanation of the sample programs by a session leader, participants are encouraged to modify, build-upon, or be inspired by, these programs to solve problems of their choosing. Participants work on projects individually, or in groups, with direct help from more experienced mentors present.

Discussion

There are several ongoing challenges with running the workshops. Among these, two that stand out are 1) supporting students who have varied interests and backgrounds and 2) supporting students beyond the workshop.

The interests and backgrounds of workshop participants vary. We collect feedback from participants after each day and debrief with instructors after each session and again after each series of workshops has concluded. Based on this process, we iterate on the curriculum and design of the workshops each time we run them. However, given the diversity of the students, the overarching question remains: What are we evaluating? Are we supporting "conversational programming"? "End user programming"? Have we expanded the pool of people who are critically informed about data science? Do we inspire a more diverse set of people to pursue a programming/data science career?

We have witnessed a variety of successful individual outcomes—UW students using the skills learned in the workshops for course projects and theses, a Seattle city program manager using open civic data to inform his department's policymaking. In addition, many of our participants have returned as mentors for subsequent instances of the workshop.

Another ongoing challenge concerns sustained engagement after the workshop. We believe that the introduction to programming and data science offered in the workshop has real utility. Yet, we want to support students who are inclined to build on CDSW to be able to do so. We use open online data throughout the workshop in part because it enables sustained engagement beyond the workshop. We maintain mailing lists, and encourage students to start meetup groups, in order to foster a community of practice around community data science.

We believe that our approach has value, and are interested in sharing the lessons we've learned with others who wish to apply the collaborative event model to new contexts. However, we acknowledge that several challenges and open questions remain, particularly around how to

more systematically evaluate the impact of our model against goals of democratising data science, and how to refine our methods to better support those goals. We look forward to discussing these issues with other researchers and practitioners working in the domain of time-bound, collaborative events.

Additional information about the Community Data Science Workshop—our motivation, model, and outcomes—is available here: https://mako.cc/academic/hill_etal-cdsw_chapter-DRAFT.pdf

About the authors

Jonathan Morgan is a design researcher with the Wikimedia Foundation. His research involves understanding the sociotechnical mechanisms by which participants in Wikimedia projects coordinate their work across time and space, and on developing tools and resources to support various models of open collaboration (including hackathons and editathons) within the the Wikimedia movement. He has organized hackathon-style workshops at CSCW 2015 and 2016 that focused on fostering community of practices around open research infrastructure.

Dharma Dailey studies how people get information during crises. She attended the first Community Data Science Workshop as a student and put what she learned into her research. She found the workshop so helpful, she stuck around to help organize more of them. She is is a PhD Candidate in Human-Centered Design and Engineering at the University of Washington

Benjamin Mako Hill is an Assistant Professor in the Department of Communication at the University of Washington. His research works to understand how the design of communication and information technologies shape fundamental social outcomes with broad theoretical and practical implications — like the decision to join a community or contribute to a public good. He is a faculty affiliate at the Berkman Center for Internet and Society and an affiliate at the Institute for Quantitative Social Science — both at Harvard University. He has also been a leader, developer, and contributor to the free and open source software community for more than a decade as part of the Debian and Ubuntu projects.

Dannon Baker
Sr. Software Engineer
Galaxy Project, Johns Hopkins University
dannon.baker@gmail.com

Keywords

Organizing, using tutorials for inclusion/inspiration, live/virtual coordination

Abstract (or, my experiences and interests as, more specifically)

I'm an event participant and organizer particularly interested in using instruction and tutorials as a vehicle by which we can improve inclusiveness and motivate productivity at hackathons, which fits into the general workshop theme of design variation. In very 'free-form' hackathons I've observed at multiple events (as both a participant and organizer) that during the early brainstorming and group coalescing stage some participants tend to get 'lost' without a particular project, task, or group and often are not able to recover from that and have a productive event. There are many ways to approach this problem, but for the Swiss German Galaxy Days 'Developer Day' this past year, we tried using short tutorials on predetermined topics as a springboard for further development. With a group of about 30 individuals, we presented short 10-15 minute tutorials followed by several hours of work stemming from the presented topic. During the day we did this rotation three times for different topics and it worked very well for ensuring that all attendees were included and had a reasonable starting point for future work. Because everyone was working from the same basic topic, it was also much easier for participants to find collaborators and seek help when needed.

I'm also interested in the workshop theme of computer-mediated communication and collaboration. More specifically, I'd like to learn how we can most successfully bridge virtual and physical events as it has happened to us in the past that individuals wanted to participate but could not physically be present at the event. We've tried to use IRC as an out-of-physical-band communication mechanism, but it seems to fall short of what we need to keep people involved as the physical-space participants tend to separate from their virtual counterparts. We also hold several purely virtual events each year with fairly good participation. When everyone is remote, coordinating with chatroom-style messaging (IRC or gitter) and organizing projects on GitHub seems to work very well, though I'd like to learn what others might be doing in this space.

Lastly, I'm interested in events bringing software developers and non-technical participants (in terms of programming) together. For the past two years we've organized a scientist and power-user focused event in conjunction with our annual hackathon. I'd like to be able to connect these interested scientists, many of whom have concrete problems, with participants at the hackathon that they can work with. This interaction, however, is difficult to fit into how we've traditionally organized our events and I'd like to learn what others have done to foster these collaborations.

Personal Background

I'm a software engineer working on the Galaxy bioinformatics platform (https://galaxyproject.org). I've been on the Galaxy team since 2010, and have been able to attend, as a participant, a varied selection of Hackathons and Codefests, some of which are listed below with rough descriptions and thoughts. I'm also the primary organizer for the yearly Galaxy Community Conference Hackathon, which is described last.

OpenBio Codefest
http://www.open-bio.org/wiki/Codefest_2012
I've been able to participate in this event twice, in 2010 and 2012. This is a collaborative, non-competitive developer driven event. Topics and groups were formed upon arrival. The structure allowed a lot of flexibility for individuals to find groups and collaborate however they saw fit. This type of 'free-form' hackathon works very well for (among other benefits) allowing people to work on things that they care personally about, but otherwise may not have been able to devote the time or resources to.

Netherlands Bioinformatics Centre (NBIC) Hackathon
https://wiki.nbic.nl/index.php/NBIC_Galaxy_Hackathon_project
This NBIC Hackathon was a two day event primarily based around trying out Agile software methods and my software project, Galaxy, was one of the three topic domains chosen as a vehicle for this. The event featured a rigid structure of sprints, and significant pre-hackathon work organizing topics and recruiting participants.

Amazon Web Services re:Invent Hackathon
https://aws.amazon.com/blogs/aws/aws-reinvent-hackathon-2013/

This was a very competitive hackathon where teams of 4-5 individuals (randomly pre-selected from conference attendees, did not meet until the day of the Hackathon) had a single day to address one of four predetermined problems.

Swiss German Galaxy Days - Developer Day
https://wiki.galaxyproject.org/Events/SG2016
For this event, several topics were selected ahead of time and 10-15 minute presentations were given, after which the entire room would work on the same problem, though in their own domain. For example, the first session involved writing installation recipes for bioinformatics tools, and many people were able to wrap and contribute new tools to BioConda (https://bioconda.github.io/) during the event.

ELIXIR Hackathon @ NETTAB 2016
https://www.elixir-europe.org/events/elixir-bioinformatics-hackathon-nettab-2016
This event split individuals into two groups, and my group attempted to organize something similar to what I had experienced at the Swiss German Galaxy Day. We used a single tutorial as a starting point, and it successfully motivated group and individual projects for the rest of the day.

Galaxy Community Conference Hackathon (2014 onward)
https://wiki.galaxyproject.org/Events/GCC2014/Hackathon
http://gcc2015.tsl.ac.uk/hackathon/
https://gcc2016.iu.edu/hacks/
I'm the primary organizer of the Galaxy Community Conference Hackathon. This is a two day event we have held each year since 2014 as a part of the Galaxy Community Conference, and it is modeled most closely after the OpenBio Codefest described above, in that it is fairly free-form and not competitive. We attempt to solicit (and seed) ideas before the event, posting them to Trello, GitHub, or the like. The Hackathon starts with a full round of introductions, and then we generally have a break for people to mingle and self-organize into groups based on their interests. So far, we have had 40-50 participants each year, for the past 3 years, and post-hackathon feedback has been quite positive. Starting two years ago, we have also organized a conjoined "Data Wrangling Hackathon" for data scientists and Galaxy power users more interested in building best practice pipelines and other non-code artifacts, and I'd like to try to encourage more collaboration between the two events moving forward.

A Typology of Hackathon Events

Margaret Drouhard*, Anissa Tanweer†, Brittany Fiore-Gartland‡
*Department of Human Centered Design & Engineering, University of Washington, Seattle, WA
Email: mdrouhard@acm.org
†Department of Communication, University of Washington, Seattle, WA
Email: tanweer@uw.edu
‡eScience Institute, University of Washington, Seattle, WA
Email: fioreb@uw.edu

Abstract—In this work we present a preliminary, exploratory typology of hackathons and similar events that has emerged from ethnographic observations. The categories and dimensions articulated here are not intended to be exhaustive or mutually exclusive descriptions, but rather to serve as lenses through which to consider various aims, strategies, and outcomes of hackathons. We describe the characteristics of three types of hackathons we have observed: communal, contributive, and catalytic. We also reference some of the observations that have led us to draw these distinctions. Finally, we consider some of the commonalities we have observed across distinct events as our understanding of hackathons continues to evolve.

Index Terms—hackathon, collaboration, ethnography

1 Introduction

Hackathons [1] have become an increasingly popular form of collaborative work for many different communities, including educators, academic collaborations, civic institutions, and non-profit organizations, among others. Hackathons have historically served as a mode of engagement across software development communities and the technology industry, but time-bound, collocated collaborative events have more recently become an emergent practice and space of inquiry in a much broader spectrum of communities. In this paper, we seek to explore the diversity of these intensive collaboration events, as well as identify commonalities among many distinct events.

As hackathons develop into an increasingly prevalent mode of short-term collaboration, researchers have begun to explore how such events serve the needs of their respective participants and organizers. Some scientific communities have designed hackathons with targeted software development or data analysis objectives, such as the "anabolic" hackathons described by Busby et al. [2]. These hackathons bring together participants who have specific skills with the aim of collaborating in the creation of prototypes to address gaps between existing public databases and tools available to explore them. The authors report that meaningful prototypes were produced during the event, and participants expressed interest in continuing to develop the tools. In our observations, we have also found that hackathons organized for communities with a particular data analysis need often include more targeted software development projects.

Strengthened social or community ties have also been identified as key outcomes of many hackathons. Nandi and Mandernach conducted observations and studied survey feedback and digital traces from an academic hackathon. They describe different models for collaboration across teams, how the hackathon served as an informal learning platform for skills not frequently taught in university classrooms, and how the event provided opportunities to develop communication skills and network with industry sponsors [3]. Trainer et al. outline insights from multiple observations of scientific hackathons. In particular, they consider the trade-offs between social and technical outcomes at hackathon events. Their observations reveal that across distinct communities and types of events, the strategies for achieving technical ends vary, but the use of collocation to strengthen social or community ties is a commonality [4]. Möller et al. also identify community-building as a core outcome of hackathons. They consider hackathons in computational biology as a means to address limitations that traditional academic conferences impose on opportunities for collaboration. They emphasize that the collaborations and stronger community ties that emerge from hackathons and similar events are often greater than the immediate tangible outcomes [5]. In our Typology of Hackathon Events, we do not include networking, strengthening of social ties, or community-building as characteristic of different types of hackathons. However, this choice is not intended to imply that we see these social components of hackathons as unimportant or uninteresting. Rather, we view them as core to any hackathon model. Our current typology focuses on distinctions among different categories of hackathons, so commonalities have been excluded.

Beyond development of specific tools and strengthening of social ties, hackathons have been proposed as a means of civic participation [6], [7]. Many municipal government agencies and other civic institutions report positive outcomes from such events, but Irani presents a more

1. We acknowledge that the word *hackathon* is not the preferred term for all such events, and is avoided by some communities and individuals [1]. We use the term "hackathon" not as a definitive term for the phenomenon, but as an emic label that is commonly used by many–but certainly not all– of our study participants.

critical perspective [8]. Irani's analysis is based on ethnographic fieldwork at a design studio in Delhi including an "entrepreneurial" civic hackathon. The hackathon in this instance was billed as a tool for civic participation, yet Irani expresses frustration that the focus was constantly on "the demo," and there was no time for "real footwork" or engaging with the communities of affected individuals. Irani laments that although the hackathon brought people together around an issue of concern, ultimately it "could not accommodate those for whom it claimed to care" [8]. Our observations have also encompassed civic hackathons, some of which have formed part of larger initiatives with ongoing development. Like Irani, we recognized that "real footwork" beyond tool-building is often fundamental to the adoption of these projects. We observe that political engagement and cross-sector collaboration has been integral to the success of projects that have lived beyond hackathon events.

In the remainder of this work, we outline our preliminary typology for hackathon events, referencing some of the events we have observed. Collectively, our team of ethnographers has spent more than 100 hours observing hackathons. Events observed include domain-specific academic "hackweeks," "hacks" for social good, civic hackathons, and hackathons for distributed research communities. We anticipate that this typology might be used as a heuristic to organize conversations around hackathon structure and aims.

2 A Typology for Hackathon Events

We present the typology shown in Table 1 as an exploratory, preliminary typology of the types of events our ethnographic team has observed. We believe the development of such a typology may facilitate the development and articulation of new lines of inquiry, research questions, and theory. The typology we present here is based in our own fieldwork observing hackathons and hack-related events and practice, and the distinctions we make between different forms of hackathons arise from our direct observations. These observations have centered around hackathons and related events that have emerged in the overlapping spaces and networks of tech industries, government, civic and social good organizations, and academia. As many of these stakeholders import aspects of software engineering culture like hackathons or sprints, these ideas are adopted and adapted.

These varying contexts have expanded our notion of the hackathon and provided fertile ground for understanding and categorizing the approaches, motivations, and characteristics of a wide-ranging collection of events and practices that are often lumped together under the hack label. It is precisely because we are situated at the intersection of these different stakeholders and sectors that we can see the multiple modes of engagement and the different motivations and expectations that shape these events and practice. Given the grounded, emergent nature of our analysis, these categories are not meant to be interpreted as mutually exclusive or exhaustive. In other words, some events may have elements of more than one category, other models of hackathons may exist that we have not observed, and new configurations may form in this quickly evolving space. Moreover, we recognize a significant amount of overlap among various types of hackathons, so we have focused our typology on the dimensions in which we observed significant distinctions.

2.1 The Communal Hackathon

Some of the hackathons we've observed were convened for the purpose of developing resources, infrastructure, practices, or culture for a particular community, and qualifications for participation included membership in that community. This community may be defined by a shared focus on a particular domain of knowledge, a shared investment in developing a particular software tool or infrastructure, or a shared interest in working with a specific type of data, language, or library. In each case, these areas of methodological common ground served as the hackathon's seed, or the given object around which the event is organized, and participants work on the challenge of advancing capacity to exploit their shared methodological approach. An important motivation for participation seems to be the opportunity to engage in professional development, both for individual participants' careers and the continuing development of professional communities. Because much of the emphasis is on developing the capacities of the community, participation is usually collaborative in nature, with people engage in a didactic style of work, in which participants are both teaching and learning from one another. We find that across the various kinds of hackathons we've observed, although the events themselves are indeed time-constrained, there is always the hope or expectation of something durable living beyond the event. In the case of communal hackathons, this desired continuity is often the community itself and its expanding methodological capacities.

As prime examples of communal hackathons, in recent years we've observed the phenomenon of "hack weeks," in which academic researchers within a specified domain of knowledge from institutions all over the world converge on a particular place for several days at a time to collaborate. These events often include tutorial sessions for relevant tools or analysis techniques as well as open time for project teams to work together. Projects may be proposed in advance or at the event itself, and they are often focused on exploring datasets of interest or on building out open source software that supports the community's analysis needs. Participants in these types of events have reported that they intend to share the tools and strategies learned from the hack week with their local communities, and some have even gone on to organize other hack events.

2.2 The Contributive Hackathon

Another kind of hackathon that we've observed is convened for the purpose of advancing the aims of a larger, pre-existing project by breaking that work into discrete, modular tasks that can be completed in short, intensive work sessions. One important qualification for participation is concern for the project or issue the hackathon is intended to address, and an important motivation for participation seems to be a desire for impact. These events are often geared toward making up for some deficit in the current state of the project, and the challenge is to complete as

TABLE 1
Typology of Hackathon Events

Keyword	Communal	Contributive	Catalytic
Primary Purpose	Developing resources, infrastructure, practice, and culture for community	Contributing to larger effort by focusing on discrete, modular tasks	Demonstrating utility of dataset, technology, or approach
Challenge	Advancing methodological approaches in the community	Completing as much work as possible in a fixed period of time	Generating an idea both novel and tractable
Seed	Specialized methods	Defined task	Articulated challenge
Motivations for participation	Professional development	Impact	Recognition
Mode of participation	Collaborate	Execute	Innovate
Style of Work Environment	Didactic	Autonomous	Competitive
Qualification for participation	Membership in community	Concern for the project/issue	Skills or an idea
Continuity	Of community	Of project	Of ideas

much work as possible in a short amount of time. In contributive hackathons, organizers often provide specific instructions for tasks that serve as the seed for the work completed during the hackathon. As such, participants' style of work is to execute those assigned tasks, and their mode of participation is often autonomous; i.e., even though they are collocated, they are often working on individual tasks. In the case of contributive hackathons, the desired continuity is for the larger ongoing project to which the hackathon contributes.

As an example, we have observed "mapathon" events organized by researchers and community members interested in mapping morphological features that are largely missing from open source mapping platforms. After developing software tools to support simple mapping of these features, researchers worked with local mapping communities to convene a number of single-day mapathons. In mapathon events, individuals are tasked with mapping specified morphological features within a narrowly defined and focused geographical area. Without that focused effort, the desired morphological data would have been added to the map much more slowly, inconsistently, and incompletely. The work of mapathons supports larger efforts of open source mapping communities by filling gaps in the network of features that have already been mapped. Organizers report that these events are particularly successful when aimed at improving representation or otherwise supporting the needs of vulnerable communities. Concern for the impact on these communities appears to be a motivating factor for participating in such events.

2.3 The Catalytic Hackathon

In contrast to the contributive hackathon that feeds into a larger pre-existing project, we've also observed hackathons that are convened for the purpose of demonstrating the use of a dataset, technology, or approach in a way that could spark the inception of a new project, application or idea. As such, the challenge put to participants is to come up with something both novel and tractable. Toward this end, as part of the qualifications for participation, individuals are expected to bring to the table skills or ideas that can be applied to this pursuit of novelty. The seed provided by organizers is often in the form of an articulated challenge that includes well-defined requirements and constraints. Frequently, organizers offer some recognition to successful project teams in the form of prize money, media attention, computational resources, or further networking opportunities as motivation for participation. In keeping with this incentive system, the style of work in these hackathons is often competitive, and participants are expected to participate through innovation. In the case of these catalytic hackathons, there seems to be a desire for continuity of the novel ideas that are generated in the course of the event.

City hackathons often exemplify the catalytic hackathon model. In these events, participants are challenged to work with a particular municipal dataset (or datasets) to discover new, pro-social, civically-oriented applications of that data. One of the ongoing projects we've been following originated in this way, winning free cloud-computing services to encourage the continued development of the project following the event. Academic hackathons also often follow this model, tasking participants with the demonstration of novel applications and offering them both prizes and mentorship from industry professionals. The entrepreneurial style of catalytic hackathons may be viewed as means to encourage participants to efficiently prototype and also find ways of marketing their ideas, allowing the ideas to continue being developed.

3 Conclusion

Our observations of hackathons and similar events have allowed us to explore these dimensions and identify unique aspects of various events. However, we have also observed several characteristics that cut across the boundaries we have defined. In general, hackathons appear to interest communities because of the opportunity to collocate and work intensively on projects or tasks in a way that is not supported by their day-to-day work environments. Additionally, as other researchers have noted, one of the most significant outcomes at most events is a strengthening of community or social bonds. As we continue to observe and participate in hackathon events, we are expanding our analysis of these commonalities as well as the distinctions among different kinds of hackathons. We anticipate the evolution of this preliminary typology as our understandings of hackathon events deepens.

4 AUTHORS

The authors are interested in participating in the Hackathon Workshop at CSCW to further explore the dimensions of hackathons and similar events, drawing from perspectives of other researchers studying hackathons and hackathon organizers.

Margaret (Meg) Drouhard is a PhD student in the Department of Human-Centered Design and Engineering (HCDE) at the University of Washington (UW). She is advised by Dr. Cecilia Aragon and works in the Human-Centered Data Science Lab. Meg's primary research interest is visual analytics, and she is most interested in opening up data exploration to non-data scientists through visualization. Meg also works with the ethnographic team and Data Science Studies Working Group as part of the eScience Institute at UW, studying the emerging practice and culture of data-intensive scientific research. As part of her ethnographic work, she has observed several academic and community-led hackathons.

Anissa Tanweer is a PhD student in the Department of Communication at the University of Washington and a research assistant in the Human-Centered Data Science Lab. She is interested in the ways people organize with and around data, and how the use of increasingly large, heterogeneous datasets is transforming the way we construct knowledge and make decisions.

Brittany Fiore-Gartland is a data science ethnographer studying the sociocultural implications of data-intensive science and how data-intensive technologies are reshaping how people work and organize. Her research focuses on cross-sector and interdisciplinary data science collaborations; emerging pedagogical models; and bringing a human-centered, sociotechnical perspective to data science practice. She received her Ph.D. in Communication from the University of Washington and an M.A. in Sociocultural Anthropology from Columbia University. As an interdisciplinary researcher she is most at home somewhere between Science and Technology Studies (STS), Media Studies, and Organizational Communication.

REFERENCES

[1] A. Decker, K. Eiselt, and K. Voll, "Understanding and Improving the Culture of Hackathons: Think Global Hack Local," in *Frontiers in Education Conference (FIE), 2015. 32614 2015. IEEE*. IEEE, 2015, pp. 1–8.

[2] B. Busby, A. Matthew Lesko *et al.*, "Closing gaps between open software and public data in a hackathon setting: User-centered software prototyping," *F1000Research*, vol. 5, 2016.

[3] A. Nandi and M. Mandernach, "Hackathons as an Informal Learning Platform," in *Proceedings of the 47th ACM Technical Symposium on Computing Science Education*. ACM, 2016, pp. 346–351.

[4] E. H. Trainer, A. Kalyanasundaram, C. Chaihirunkarn, and J. D. Herbsleb, "How to Hackathon: Socio-technical Tradeoffs in Brief, Intensive Collocation," in *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, 2016, pp. 1118–1130.

[5] S. Möller, E. Afgan, M. Banck, R. J. Bonnal, T. Booth, J. Chilton, P. J. Cock, M. Gumbel, N. Harris, R. Holland *et al.*, "Community-driven development for computational biology at Sprints, Hackathons and Codefests," *BMC bioinformatics*, vol. 15, no. 14, p. 1, 2014.

[6] P. Johnson and P. Robinson, "Civic Hackathons: Innovation, Procurement, or Civic Engagement?" *Review of Policy Research*, vol. 31, no. 4, pp. 349–357, 2014.

[7] S. J. Carr and A. Lassiter, *Big Data, Small Apps: Premises and Products of the Civic Hackathon*. Cham: Springer International Publishing, 2017, pp. 543–559. [Online]. Available: http://dx.doi.org/10.1007/978-3-319-40902-3_29

[8] L. Irani, "Hackathons and the Making of Entrepreneurial Citizenship," *Science, Technology & Human Values*, vol. 40, no. 5, pp. 799–824, 2015.

Challenges Incorporating Community Feedback at Recurring Civic Hackathons

Eureka Foong
Northwestern University
Evanston, IL 60208, USA
eureka@u.northwestern.edu

Elizabeth Gerber
Northwestern University
Evanston, IL 60208, USA
egerber@northwestern.edu

Abstract
Civic organizers are increasingly appropriating hackathons to gather volunteer programmers, designers, and subject matter experts to develop technical solutions to social issues. Despite their increasing frequency we know little about how hackathon organizers support these events. We conducted a 6-week participant observation of a weekly civic hackathon in the Midwest. Our analysis suggests that organizers face three challenges with user research: 1) balancing user research with development, 2) testing early concepts with the target population, and 3) communicating research insights with new hackathon participants. We present opportunities for sociotechnical researchers and hackathon organizers to improve the way civic hackers incorporate feedback from the communities they serve.

Author Keywords
Hackathons, civic, collaboration, feedback, community, participatory design, user research

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction and Background

Throughout the globe, programmers, designers, and domain experts join hackathons, short and intensive events that produce technical solutions to challenges in areas such as computational biology and computer science [7]. More recently, programmers, designers, and domain experts join hackathons to address "social conditions and their consequences" [2], tackling issues such as wildlife conservation and the prevention of bacterial outbreaks.

Civic hackathons have emerged as one way to develop technology outside of formal design environments [2]. These time-bounded events leverage the talent of programmers, designers, and subject matter experts to develop applications that directly improve civic life [2]. While these events help participants recognize their role in shaping governance [6], multiple HCI scholars are doubtful of their technological productivity [2-4,6]. The fleeting nature of hackathons makes it unlikely that innovative, deployable solutions will be developed [2]. In fact, Trainer et al. [9] stress the need to develop tools that support continued work after the hackathon.

Unlike short-term hackathons that may lack long-term impact, recurring civic hackathons allow teams to iteratively test, develop, and launch their solutions in the community. For example, ad-hoc "brigades" of civic hackers within the organization Code for America meet regularly to work on new or existing projects [1]. Despite the adoption of this model across several U.S. cities [1], we still know little about how "hacking" should be supported [5]. This is partly because extant research examines the value of these events in generating civic participation, rather than in producing feasible and sustainable solutions [2-4,6]. As a result, little research has sought to improve the productivity of civic hackathons.

In this paper, we focus on one area of the civic hackathon that warrants improvement: developing a shared understanding of the design challenge [4]. Participants at short-term hackathons struggle to develop a rich understanding of their end users and the design challenge, and scaffolding the design process may help participants overcome these difficulties under strict time constraints [4,8]. Another study suggests that participants who are also end users of the solution provide valuable feedback that greatly improves the quality of proposed solutions [9]. Nonetheless, it is unclear how these need-finding and evaluation activities unfold and can be supported in a recurring civic hackathon.

The research question driving the current study is: how does an ad-hoc design team at a recurring civic hackathon understand its users and the complexity of the challenge to develop a technological solution? Through a 6-week field study at a weekly "hack night" in the Midwest, we learned civic hackers struggle to simultaneously conduct user research and develop technical solutions. To overcome some this challenge, we propose hackathon organizers invest more effort in incorporating community feedback at these events. We call on sociotechnical researchers to study how these efforts could influence the productivity and impact of civic hackathons.

A Case Study of a Recurring Civic Hackathon

From January to March 2016, we conducted a 6-week ethnographic study of a weekly civic hackathon in a large city in the Midwest. To preserve the confidentiality of our participants, we have used pseudonyms to represent the informants, organizations, and locations involved in our study. The hackathon is scheduled every week from 6:00 to 10:00 p.m. at the office of an online software company. The majority of the 40-100 attendees are professionals from the technology industry, journalists, freelancers, and local graduate students, who were attracted to the event for its career and learning opportunities.

During hack nights, participants join breakout groups that work on projects or learning groups that provide support for developing technical skills. One of the authors joined the Justice in the City breakout group led by Kelly, a nonprofit attorney. The group's mission was to develop a technological solution to recidivism by removing information barriers that lead the formerly incarcerated to violate parole. The membership of the group fluctuated from week to week, with only a couple of individuals that could be considered core members. Group members have included geographic information system (GIS) analysts, programmers, journalists, user interface designers, and writers. The size of the group varied between 3 and 7 people.

Challenges Balancing User Research with Development

Throughout the study, Justice in the City devoted time to understanding problems facing their target community, but struggled to simultaneously develop and test technical solutions. Early on, Kelly invited individuals who have worked with the formerly incarcerated to speak to the team. These discussions directed us away from less feasible ideas; for example, we quickly learned that a mobile application would not be ideal for providing information as many parolees do not have access to a personal smartphone. In addition, these discussions helped the team understand larger systemic problems that lead to recidivism. Nonetheless, these new insights also prevented us from deciding on suitable technical solutions. One member expressed concern that the team had not started development despite working on the project for a few weeks.

When solutions were explored, the group failed to test early concepts with target users. After learning that websites for parolees contained outdated, difficult-to-use resources, the team attempted to create an online guide for re-entering society based on a guide developed by a local organization. We spent one meeting reading an existing guide to familiarize ourselves with the resources that should be provided to parolees. Although we listed ways to improve the organization of this content, we were unable to test these solution ideas directly with the target population. This was partly due to the difficulty of arranging regular meetings with recent parolees and providing transport to the hackathon venue.

Because not all members were present at each meeting, the team also struggled to communicate research findings across hack nights and especially to new members. Only core members had attended events outside of hack night (i.e., a job fair and a support group meeting) to understand the target population. Other team members, including the author, were assigned remote tasks, such as researching websites for recent parolees. While the author helped

the group start a virtual task and ideas list to organize these research findings, new members still struggled to tap into this shared understanding of the problem space and productively contribute to discussions about solution ideas. In order to accommodate this, the leader would summarize key insights from previous hack nights for new members, which was time consuming.

Discussion and Limitations

While the civic hackathon we attended provided ample technical resources, our study revealed difficulties hackers face when learning about the communities they serve. To address these difficulties, hackathon organizers should explore ways to help teams incorporate feedback into the development process. For instance, organizers could provide themed events to encourage teams to alternate between developing technical solutions and conducting user research. To reduce the burden on individual teams, organizers can establish relations with subject matter experts in the community and invite these individuals to provide feedback at hack events.

Sociotechnical researchers should also investigate the use of technology to support civic hackers; for example, online crowdwork platforms could be used to test early concepts with a large number of users. Lastly, researchers should study how technology can improve shared understanding of insights from user research. Because our study is focused on one civic hackathon team over a limited period of time, further research is needed to understand the broader spectrum of challenges civic hackers face in developing technology for their communities.

Conclusion

Even though investing time in user research helps civic hackers develop more feasible solutions, it is challenging for teams to simultaneously develop technical solutions, test concepts early with target users, and communicate research insights with new team members. We call on hackathon organizers to become more committed to helping teams incorporate community feedback into their projects by increasing access to subject matter experts from the community. We also call on sociotechnical researchers to study solutions for crowdsourcing feedback online and sharing research insights. At CSCW 2017, we are excited to discuss these and other opportunities for increasing community participation at recurring civic hackathons.

About the Authors

Eureka Foong is a PhD student and Segal Design Cluster Fellow at Northwestern University studying crowdsourcing applications that support design education. She has spoken at TEDx about user research and problem solving at civic hackathons.

Elizabeth Gerber, PhD, is the Breed Junior Chair of Design at Northwestern University and Faculty Founder of Design for America, a nationwide network of student design teams. Her current research focuses on understanding the work of social innovators in online and offline communities.

Acknowledgements

We would like to thank the organizers and informants in our study, Professor Gary Fine in the Department of Sociology at Northwestern University, as well as our colleagues in the Delta Lab for their helpful feedback and NSF Awards IIS 1217225 and 1530837.

References

1. Code for America. 2016. List of all Brigades - Code for America. *Code for America*. Retrieved December 15, 2016 from https://www.codeforamerica.org/join-us/volunteer-with-us/list-of-all-brigades.
2. Carl DiSalvo, Melissa Gregg, and Thomas Lodato. 2014. Building belonging. *Interactions* 21, 4: 58–61. https://doi.org/10.1145/2628685
3. Melissa Gregg. 2015. Hack for good: Speculative labour, app development and the burden of austerity. *The Fibreculture Journal* 186, 25: 183–201. http://doi.org/10.15307/fcj.25.186.2015
4. Lilly Irani. 2015. Hackathons and the making on entrepreneurial citizenship. *Science, Technology & Human Values* 40, 5: 799–824. http://doi.org/10.1177/0162243915578486
5. Peter Johnson and Pamela Robinson. 2014. Civic Hackathons: Innovation, Procurement, or Civic Engagement? *Review of Policy Research* 31, 4: 349–357. http://doi.org/10.1111/ropr.12074
6. Thomas James Lodato and Carl DiSalvo. 2016. Issue-oriented hackathons as material participation. *New Media & Society* 18, 4: 539–557. http://doi.org/10.1177/1461444816629467
7. Steffen Möller, Enis Afgan, Michael Banck, et al. 2014. Community-driven development for computational biology at Sprints, Hackathons and Codefests. *BMC Bioinformatics* 15, Suppl 14: S7. http://doi.org/10.1186/1471-2105-15-S14-S7
8. Emily Porter, Christopher Bopp, Elizabeth Gerber, and Amy Voida. 2017 (forthcoming). Reappropriating hackathons: The production work of the chi4good day of service. *Proceedings of the 2017 Conference on Human Factors in Computing Systems*.
9. Erik H Trainer, Arun Kalyanasundaram, Chalalai Chaihirunkarn, and James D Herbsleb. 2016. How to hackathon: Socio-technical tradeoffs in brief, intensive collocation. *Proceedings of the 2016 Conference on Computer-Supported Cooperative Work*: 1116–1128. http://doi.org/10.1145/2818048.2819946

Insights from Hackathons in Action: A Gold Mine for studying People's Moods & Behaviors

Krithika Jagannath
PhD Student in Informatics
University of California, Irvine
kjaganna@uci.edu

ABSTRACT

Hackathons, known to have been introduced in the late 90s [1], manifest in a variety of ways and broadly comprise creative and/or problem-solving activities. An example of one such variant is stitch-fest [2], a make-a-thon where makers, artists, and technologists converge to produce outcomes by designing or "making". While the formats of the different kinds of hackathons vary, there is a shared understanding within the community in terms of the temporal, physical or virtual, and social characteristics that go into categorizing an event as a hackathon. This paper is focused on a human-centered application for hackathons – through which it we can deepen our understanding of people's moods and behaviors in collaborative settings. Of significant interest is the impact of people's moods and behaviors on the outcomes for both the community and the hackathon(s). Hackathons and technology-based start-ups are similar in a few obvious ways but the former still represents an exaggeratedly scaled-up version in terms of time and collaborative characteristics. In light of the dynamic social interactions on a formal and informal level, hackathons offer a rich landscape for studying people's moods and behaviors in collaborative settings. Sensor-based technologies in which sensors are used in the physical environment and in unobtrusive wearables can help us gather rich data that represents participants' stress levels, moods, and behaviors. The same data will allow us to closely examine the design and privacy implications of sensor-based technologies. More importantly, the data can help us assess long-term and short-term impact to individuals, communities, and to hackathons. Eventually, we might use this data to drive design enhancements and tools' development for future hackathons.

INTRODUCTION

In this position paper, the idea proposed is a human-centered application for hackathons through which we can further understand how dynamic time-bound environments impact people's moods and behaviors. Doing so can reveal insights into how future hackathons should be designed (physically and virtually) to better support individuals, communities, and event organizers.

APPLICATIONS

Hackathons have widespread applications – as learning environments [e.g., 3], for product development à la technology start-ups, and for networking and hiring even [4] to name a few examples. The hackathon phenomenon usually lacks an absolute rigid structure but occurs within some constraints. For example, let us consider a hackathon in which participants are required to come up a solution for a stated problem as teams of four within 48 hours (constraints), but acceptable solutions could involve design, an app, a program, and a range of possibilities (i.e. no structure imposed). Similarly, hackathons usually are time-bounded (constraints) but do not interfere with collaborations that participants can build on beyond the scope of the hackathons (i.e. no structure imposed). Hackathons might require people to self-organize as teams

(constraint) but some may allow collaborations across teams (i.e. no structure imposed). Given the potential tensions between opposite extremes of constraints and lack of a rigid structure in hackathons, how exactly do people's moods and behaviors unfold? How do the sizes and durations of hackathons influence people's reactions to the demands of constraints? How do the differences in skill levels both within and across teams that are manifested as competition or cooperation impact people's moods and social behaviors? How can people's affective states [5], stress-levels (both positive and negative arousal), moods, and behaviors as data to infer what kinds of challenges and barriers exist in hackathon environments?

Emerging sensor-based technologies in which sensors are embedded in the physical environment and present in unobtrusive wearables can help us gather rich information about individuals' stress-levels, arousal states, affective and mood related data, and individual and social behaviors. This data can reveal insights into challenges and barriers that individuals experience in hackathons and also help us understand the nuances of time-sensitive collaborations. It can also prompt organizers of hackathons to make changes in the design of hackathons

PRACTICAL SUPPORT FOR EVENT ORGANIZERS

Participants' moods and behaviors in the context of hackathons can offer many insights to event organizers. To name a few:

1. Barriers and challenges that participants experience at various stages of the hackathon,
2. How hackathon communities develop norms around collaborative work,
3. People's affective stages during various phases before, during, and after hackathons.

Similar to the ongoing efforts to challenge the "one size fits all" notion in education, this approach could lead to positive changes in one or more ways – in how hackathons are scheduled of events, in more universal formats to support needs of different kinds of individuals, and so on.

SHORT-TERM AND LONG-TERM OUTCOMES

Enabling individuals monitor and understand how their stress levels, affective states and moods might change over time during their engagement in the hackathons can help them manage their own professional and personal goals. An increased self-awareness can have short-term and long-term outcomes for the ways in which people engage in the time-bound events as well as ongoing community interactions. While this remains primarily a human-centric focus for the outcomes, the success of hackathons depends on various behavioral aspects of its participants. Therefore, knowing more about what support structures enhance or limit people's engagement in hackathons can have broader implications for the field itself.

SUMMARY, MOTIVATION, AND BACKGROUND

In closing, the idea proposed in this paper is about an extended application of hackathons. Sensing technologies can help us gather information about people's stress levels, moods, and behaviors during their participation in hackathons. From this data, we could expect to gain insights into needs of participants, hackathon communities, and how physical or virtual events might need to be organized in the future. In addition, we might learn more about how norms within hackathon communities emerge and how changes could be brought about to enhance the

experience for those individuals within hackathon communities and the organizers themselves. In the process gathering data, we could derive added benefits of improvements in design or privacy aspects of sensor-based technologies through a deeper understanding of people's notions of such technologies. While a hackathon was deployed in a study [6] around personal informatics, studies in hackathons that focus on enabling people to learn more about their mental and emotional states have not been fully explored.

As an ardent programmer and maker, I am interested in how human-centered technologies could be designed to enable people to collaborate and produce creative work that is personally relevant and meaningful to them. Hackathons are particularly interesting because of the dynamism brought into play when people collaborate to create, learn, and solve problems. Human-Computer Interaction or HCI is my chosen field of research. My background in Computer Science Engineering and the Learning Sciences enables me to integrate diverse perspectives in my work. Broadly, my interests are areas of Computer Supported Collaborative Work (CSCW) and Computer Supported Collaborative Learning (CSCL).

REFERENCES

[1] Briscoe, G. (2014). Digital innovation: The hackathon phenomenon.

[2] Richard, G. T., Kafai, Y. B., Adleberg, B., & Telhan, O. (2015, February). StitchFest: Diversifying a College Hackathon to broaden participation and perceptions in computing. In Proceedings of the 46th ACM Technical Symposium on Computer Science Education (pp. 114-119). ACM.

[3] Nandi, A., & Mandernach, M. (2016, February). Hackathons as an Informal Learning Platform. In *Proceedings of the 47th ACM Technical Symposium on Computing Science Education* (pp. 346-351). ACM.

[4] Leckart, S., (2015) The Hackathon Fast Track, from Campus to Silicon Valley, Article from the New York Times dated April 6, 2015. Retrieved from the web on December 13,2016 from http://www.nytimes.com/2015/04/12/education/edlife/the-hackathon-fast-track-from-campus-to-silicon-valley.html

[5] Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. Development and psychopathology, 17(03), 715-734.

[6] Li, I., Froehlich, J., Larsen, J. E., Grevet, C., & Ramirez, E. (2013, April). Personal informatics in the wild: hacking habits for health & happiness. In *CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 3179-3182). ACM.

Knowledge Sharing in Time Bounded Collaborative Events – Connecting Theory and Practice

Arun Kalyanasundaram
Carnegie Mellon University
Pittsburgh, PA 15213, USA
arunkaly@cs.cmu.edu

James D. Herbsleb
Carnegie Mellon University
Pittsburgh, PA 15213, USA
jdh@cs.cmu.edu

Abstract
Knowledge is one the most useful entities exchanged among members in a collaborative activity. The time-bounded nature of hackathon like events makes it important to optimize the process of knowledge sharing to achieve the desired goals. We use a model of knowledge creation from organizational behavior literature to formalize the notion of knowledge sharing in the context of time-bounded collaborative events. We discuss practical implications of the model, present open questions and ideas for further discussions at the workshop.

Author Keywords
Time-bounded collaborative events, Knowledge sharing, Event design, Hackathons, Tacit knowledge

ACM Classification Keywords
H.5.3. Information interfaces and presentation (e.g., HCI): Group and Organization Interface–Computer-supported cooperative work

Introduction
Teams need to share knowledge among its members in an efficient and effective way in order to perform their tasks [1]. Depending on the context, the knowledge required to perform a task may be quite diverse. For

example, software development often requires knowledge of the domain, technical know-how, project practices and policies, programming conventions and knowledge of who knows what, among others. The time bounded nature of events like hackathons makes it even more important to optimize the process of knowledge sharing. We believe that depending on the goals of a hackathon, it may be appropriate to facilitate different forms of knowledge sharing. For example if the goal of a hackathon is learning new tools then knowledge transfer through hands-on activities and inter-personal interactions may be more appropriate, whereas if the goal is to complete inter-related tasks then group meetings and sharing knowledge across teams might be more appropriate. Therefore, a theoretical framework to better understand the process of knowledge sharing in hackathon like events is necessary. Our goal is to build on existing theories related to knowledge sharing from organizational behavior literature. We are interested in both the theoretical space around hackathon process and the practical implications of our propositions for organizers of hackathon like events.

Knowledge Creation and Sharing

In this section we discuss relevant literature on how researchers have conceptualized knowledge and modeled the process of knowledge sharing. In the next section we will explore different ways these concepts and theories may be useful in the context of time bounded collaborative events.

Human knowledge is beyond just possessing information, it allows us to define, prepare, shape, and learn to solve a task or problem [2]. Polanyi [3] proposed that human knowledge manifests in two forms: a) explicit and b) tacit. Explicit knowledge is objective and rational knowledge that can be precisely articulated in words, numbers and figures. For example, how to use a printer, solution to a differential equation, etc. Tacit knowledge on the other hand is subjective, experience knowledge that is not easily articulable and in the minds of the people. For example, wine tasting, social norms and even leadership skills. Explicit knowledge is also referred to as codified or know-what, while tacit knowledge may be called noncodified or know-how. Perhaps the most important distinction is that knowledge is explicit and tacit along a continuum. This means that tacit knowledge is somewhat articulable if it leans towards the explicit side of the continuum.

Nonaka [4]–[6] proposed a theoretical framework to conceptualize the process of knowledge creation and sharing based on a spiral model of conversion of one form of knowledge to another. We briefly discuss this model and use it as a theoretical basis to formalize knowledge sharing in hackathon like events.

Figure 1 shows a continuous loop illustrating knowledge conversion from one form to another along the explicit-tacit continuum and each cell represents the process by which this transformation occurs. The first cell ‘socialization’ indicates the transfer of tacit knowledge from one person to another person happens through inter-person learning. This can also happen with observation, imitation and practice. The second process is ‘externalization’ which is conversion of tacit to explicit. One case is articulation of one’s own tacit knowledge in words, metaphors and analogies. A second case is eliciting and translating the tacit knowledge of others in a readily understandable form

(e.g. customer requirements). The third process is 'combination' – conversion from between one form of explicit knowledge to another. For example, consolidating meeting discussions into minutes, processing documents to make it more usable. The last part in this cycle is 'internalization' which is understanding or absorbing explicit knowledge in to tacit knowledge held by the individual. The theory states that knowledge is not only transformed conversion process could also lead to the creation of new knowledge at each stage and hence the model is referred to as the knowledge creation model.

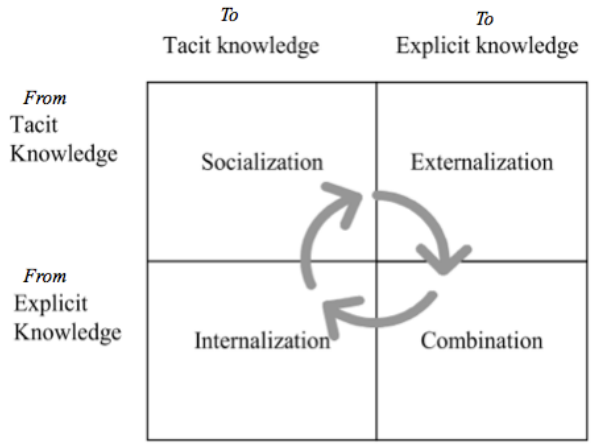

Figure 1: A Model of Knowledge Creation

Applying the Knowledge Creation Model

In this section we discuss knowledge sharing during the different stages of time bounded events using this model of knowledge creation.

Preparation Stage

The preparation stage is the period that is a week or two before the start of a time bounded event. In this period participants are unlikely to interact face to face (except e.g. people working in the same office). In our experience we observed that this stage is used to propose ideas to be worked on at the hackathon and get to know about each other. This involves absorbing explicit knowledge in the form of discussions and participant profiles on their website into tacit knowledge. This can lead to the creation of a transactive mental model where people learn about others expertise (who knows what), and possible projects they can work on. This tacit knowledge is often incomplete and may be inaccurate at times depending on the amount of explicit knowledge that is made available through ICT tools.

Participants who are proposing ideas in this stage are involved in either the conversion of tacit to explicit (externalization), or explicit to explicit knowledge (combination). This sort of externalization is similar to eliciting requirements from customers and is widely studied in the literature [7]. People also prepare demos and tutorials to be used during the event, which is accomplished through the process of combination. Since tools to accomplish combination have been designed in information systems, perhaps we could leverage those tools to assist participants in this process.

Hacking Stage

During the event itself, as we discussed earlier optimizing knowledge sharing is extremely important due to the time bounded nature of the event. As a result it is important to prioritize certain forms of

knowledge sharing over others. Hackathon like events might be conducive to share tacit knowledge because – a) participants are often collocated and b) there is emphasis on 'doing' things. Socialization might seem an ideal form of sharing tacit knowledge in these events. However, for hackathons where learning is a goal, it may not be scalable to maximize sharing tacit knowledge through inter-personal learning and socialization. At the same time learning from documents (explicit knowledge) might be time consuming. Therefore, bootcamps offer a nice tradeoff between these two forms of learning and are found to be effective forms of knowledge sharing at hackathons for learning outcomes.

Inter dependent tasks across teams might give rise to knowledge sharing across teams. While sharing explicit knowledge might seem straight forward, it may not be readily consumable by another team. Therefore, adopting the process of combination might be necessary in knowledge sharing across teams. For example, exposing certain functionalities in the code through adding wrapper functions. Sharing tacit knowledge through socialization (e.g. meetings) across teams is possible but could break the rhythm of work in certain teams. So it is important to keep track of tacit knowledge that cannot be externalized and specifically raise share those during inter team meetings.

Another important form of tacit knowledge that is often difficult to articulate requirements for software tools [7], [8]. A face to face setting allows users to easily share new use cases and show specific ways of using tools to developers as these are often difficult to articulate.

As a result emphasis must be made to identify and share tacit knowledge during the hacking stage to make use of the face to face setting. At the same time, an important but often overlooked form of knowledge sharing is combination. Discussions among team members are often informal and available in the form of meeting notes or whiteboard sketches. These are explicit knowledge but are not easily usable outside the context of the event. This is particularly important in events where discussions and tasks need to be followed up after the event is over.

Follow Through Stage

In certain types of events, participants continue working on their tasks even after the event, albeit in a distributed setting. As a result knowledge created during the event must be externalized in a form that can be easily consumable after the event. We have observed organizers encouraging this by creating a wiki page for each team and allowing participants to document relevant information in a predefined template format. Even though teams might interact via video conferencing and other ICT tools the loss of context makes it difficult to share knowledge in a distributed setting unless it is externalized and combined during the event.

Implications for Practitioners

The knowledge creation model gives us a taxonomy to understand and discuss the intricacies of knowledge sharing process during time bounded events. One of the practical implications of this analysis is to be able to prioritize certain forms of knowledge sharing during the different stages of the event depending on the desired outcomes.

Let's say the goal of an event is learning, e.g. to enable participants learn new tools, techniques and concepts of a scientific domain. In this case the preparation phase might benefit from creation of tutorials, demos, etc. This can be viewed as knowledge sharing through externalization (tacit to explicit) and combination (explicit to explicit). During the event, depending on the whether participants acquire more tacit or explicit knowledge, the emphasis on can be either on socialization (tacit to tacit) or internalization (explicit to tacit). The follow through stage of such an event might involve participants sharing their new tacit knowledge with other members of the community aka socialization. Table 1 illustrates this analysis, the X in each cell represents the form of knowledge sharing that organizers might pay attention to during each stage of a given event.

	S	**E**	**C**	**I**
Preparation		X	X	
Hacking	X			X
Follow-Through	X			

Table 1: Desired forms of Knowledge sharing during different stages of an event with 'Learning' as a goal. S: Socialization, E: Externalization, C: Combination, I: Internalization.

Similarly, let's say the goal of an event is to build finished products or complete a set of tasks collaboratively. During the preparation stage participants might want to learn about other participant's interests, skills, and discuss ideas. During the event, dependencies across tasks could make knowledge sharing across teams important. Depending on the context wither socialization (team meetings) or externalization (e.g. creating design documents that other teams can absorb (internalization)) might be beneficial. Incomplete tasks might get worked on during the follow through stage, which requires preserving work in certain forms during the event. A less favorable but inevitable form would be articulating one's own tacit knowledge during the follow through stage through externalization. This analysis is illustrated in Table 2.

	S	**E**	**C**	**I**
Preparation		X		X
Hacking	X	X	X	X
Follow-Through		X	X	

Table 2: Desired forms of Knowledge sharing during different stages of an event with 'task completion' as a goal.

Further Points of Discussions at the Workshop

An interesting application is to compare between desired and actual forms of knowledge sharing, and to what extent does the mismatch between the two influence the outcome of the event?

Our analysis is based on the premise that knowledge can be categorized into tacit and explicit, however, to what extent can we identify these in the context of hackathon like events?

One of the things we did not discuss but could be more important are the practical challenges associated with sharing knowledge. Prior work in software engineering shows that resistance to be known as an expert as a major barrier to facilitating knowledge sharing [9]. We

aim to learn about the challenges and barriers to facilitate knowledge sharing in these events from both practical experiences and established theories.

Author Biographies

Arun Kalyanasundaram is a PhD student in the Institute for Software Research at Carnegie Mellon University. His research involves studying coordination and collaboration in open-source software. He has performed ethnographic studies of hackathons to understand their socio-technical outcomes and their impact on building scientific software communities.

James Herbsleb is a Professor of Computer Science at Carnegie Mellon University, where he serves as Director of the PhD program in Societal Computing. His research interests focus on global software development, open source, and more generally on collaboration and coordination in software projects. He was recently awarded the SIGSOFT Outstanding Research Award in 2016, and previously the Alan Newell Award for Research Excellence in 2014. He has served on the PC of several conferences, including ICSE and FSE, was co-chair of CSCW 2004, and served as an associate editor of ACM Transactions on Software Engineering and Methodology.

References

[1] S. Ryan and R. V O'Connor, "Acquiring and sharing tacit knowledge in software development teams: An empirical study," *Inf. Softw. Technol.*, vol. 55, no. 9, pp. 1614–1624, 2013.

[2] G. Von Krogh, K. Ichijo, and I. Nonaka, *Enabling knowledge creation: How to unlock the mystery of tacit knowledge and release the power of innovation*. Oxford University Press on Demand, 2000.

[3] M. Polanyi, *The Tacit Dimension.* --. Doubleday, 1966.

[4] I. Nonaka, "A dynamic theory of organizational knowledge creation," *Organ. Sci.*, vol. 5, no. 1, pp. 14–37, 1994.

[5] I. Nonaka, *The knowledge-creating company*. Harvard Business Review Press, 2008.

[6] I. Nonaka and G. Von Krogh, "Perspective-tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory," *Organ. Sci.*, vol. 20, no. 3, pp. 635–652, 2009.

[7] A. H. Mohamed, "Facilitating tacit-knowledge acquisition within requirements engineering," in *Proceedings of the 10th WSEAS international conference on applied computer science (ACS'10), world scientific and engineering academy and society (WSEAS), Stevens Point, Wisconsin, USA*, 2010, pp. 27–32.

[8] J. Wan, D. Wan, W. Luo, X. Wan, and others, "Research on Explicit and Tacit Knowledge Interaction in Software Process Improvement Project," *J. Softw. Eng. Appl.*, vol. 4, no. 6, p. 335, 2011.

[9] K. C. Desouza, "Barriers to effective use of knowledge management systems in software engineering," *Commun. ACM*, vol. 46, no. 1, pp. 99–101, 2003.

Planning Inclusive, Innovative Design Sprints

Victoria Schwanda Sosik
Google, Inc.
New York, NY 10011
victoriasosik@google.com

Abstract
As part of my role as a User Experience Researcher at Google, I have participated in and led 10+ design sprints of various types. I have seen both successes and failures in these sprints, and am interested in learning how to continue to evolve this method, prevent it from becoming stale, and making sure it is an inclusive environment where all participants can contribute. I hope that conversations with workshop participants, both theoretical and applied, will help me continue to work toward these goals.

Author Keywords
Design sprints; design thinking; innovation.

Background
I'm currently a Sr. UX Researcher at Google and have been working as part of the Maps team for the past 2.5 years. In this role, I lead research around user-generated content and dining efforts, working with the product teams to conceive, plan, iterate on, and evaluate related experiences. I have been trained in Google's sprint approach[1] through our Sprintmaster Academy, and have participated in and led 10+ design sprints with Googlers, startup teams, and university students in the US and abroad. The goals of these sprints have varied widely from coming up with innovative ideas in a space where teams have been working for years,

[1] http://www.gv.com/sprint/

to creating a vision for a new product team and working towards getting the entire team on the same page.

While my experiences with sprints have been purely in industry, I also have an academic background; prior to joining Google, I completed a PhD in Information Science from Cornell University. A good portion of my time since then can be described as learning how to grow this skillset to be effective as an industry researcher. Sprints are one of my favorite ways to incorporate research into product thinking as they provide the space for a group of cross-functional team members to come together with the goal of generating new ideas and shared understandings. That being said, I am unfamiliar with the literature on sprints, hackathons, and the like, and I look forward to the ability to learn about this theoretical background from other attendees while brainstorming how to use these insights to continue evolving my use of sprints in practice.

Themes of Interest

Theoretical space of 'hackathons'
As noted above, I'm very interested in learning from others about the theoretical foundations of time-bounded events like sprints. Specifically, I'd be interested in understanding these theories so I can think more deeply through how to apply them to my work on sprints and to use them to inspire new design sprint methods (and vice-versa, what can learnings about sprints do for other types of sprints).

I've also been recently inspired to think about ways that different cultural and personal differences may be taken into consideration when planning such events. After I (along with colleagues) led workshops to teach and use design thinking methods during the inaugural AfriCHI conference, I had a discussion with a participant from Egypt about the experience. She asked me how well I thought the teams performed during the workshop in learning and applying the methods we covered. I thought many things went well, but also noted some areas where our usual process wasn't as successful as I had seen it be in the past. This sparked a conversation around ways that the process may not fit the African culture. For example, the very essence of sprints is their time-constrained nature, however I was told that in Africa people are not accustomed to being put under this pressure and therefore may need more guidance on how to allocate time. So it was suggested that instead of giving one large chunk of time for an activity, it should be broken down into its parts with each part being timed. While this is one example, I'm also interested in thinking through the types of cultures, personalities, and other participant differences that are better and worse represented in our usual processes and how we can evolve them to be more inclusive.

Design variations
Having participated in and led many sprints, I believe I can contribute much to the discussion around the practical aspects of running sprints, and stories of success and failure. I've been doing quite a bit of reflection around the sprint process having recently come off of organizing what I would consider to be my most successful and inspiring sprint to date, given that we had an explicit goal of avoiding some of our previous pitfalls. I'll summarize two aspects of the sprint that I found to be particularly effective and how that I'm interested in continuing to explore along these lines:

- The first aspect is all around helping participants shift their perspective. When I first started seeing the design sprint method being used in my teams, it was a novel way of getting a group of people together to quickly ideate, work through, and agree upon complex design decisions. However I saw the method

become a bit stale as team members were used to doing the same exercises over and over again. Dedicated to not replicating the same sprint for the nth time, a colleague and I set up to plan a sprint that was designed to introduce a provocative topic in order to shift thinking and inspire innovation. We saw that when participants no longer felt comfortable about a topic they were forced to not just fall back on old ideas, but really think through the space that felt new and interesting, even though it was really the same design space, but with an added twist. I'm interested to talk with participants about whether or not we need to continue to evolve our methods in order to keep producing the desired results, or what other techniques are out there to avoid the pitfall of running the same sprint over and over.

- The second aspect is around optimizing the attendee list for success. I have seen a tension in this space between wanting to include all the stakeholders and incorporating a lot of people's ideas, and having the "right" people to be able to more quickly converge on decisions to support the sprint's outcome. In this most recent sprint we employed a hybrid model where there was a core group of participants and a larger group, and they both played roles throughout. We did a large group kick off for 3 hours on the first morning where we generated a lot of ideas. We then broke into the small group in order to refine the ideas that came out of the brainstorm and to make quicker decisions. This model worked well to make a large group of people feel head and to secure buy in, while not compromising the ability to move quickly and make fast decisions in the smaller group. I'm interested in exploring other participation structures in order to achieve a variety of different sprint goals.

Challenges in community-building through Google Summer of Code for scientific projects

Ryan R. Curtin, principal maintainer of mlpack (`http://www.mlpack.org/`)

December 17, 2016

1 Talk abstract

For the past four years (2013–2016), the mlpack C++ machine learning library has participated in the Google Summer of Code program. The original goals of participation in GSoC was to attract new contributors from around the world, and in this way one could say that the participation was successful—in part because of GSoC, mlpack now has over 60 contributors and a fairly active community.

However, the Summer of Code program is a double-edged sword, and it exacts a heavy toll on maintainers. mlpack's somewhat unique position as one of few projects in GSoC focused on machine learning compound this toll. Last year, mlpack received 119 applications for 6 positions. There were hundreds of mailing list posts during the year from unqualified but optimistic prospective Summer of Code students seeking help; even during the off season, the mlpack IRC channel typically helps one or two people per week whose stated intentions are solely to participate in Summer of Code. Further exacerbating the difficulty, the difficult knowledge requirements to successfully contribute to mlpack (machine learning and C++ knowledge) mean that finding good candidates who will continue to be engaged with the project after the Summer of Code ends is a time-consuming and difficult proposition.

This presentation will discuss my own personal experiences with the Google Summer of Code program, which are extensible and applicable to any scientific project planning to participate in any similar program or participate in a hackathon. My focus will be on the strategies that mlpack has developed for finding high-quality candidates and keeping them engaged, but the strategies are not complete so we are left with some open questions that I won't claim to have the answer to (only some ideas and directions):

- In a time-limited environment such as a hackathon, how can a participant quickly contribute to a library full of complex algorithms?
- Are hackathons an appropriate tool to use in trying to grow the community of a scientific project?
- In what ways can we maintain a project consisting of difficult-to-understand implementations while keeping the barrier to entry low?

It might be appropriate to either pose these as questions for later discussion, or simply turn the talk at some point into an interactive discussion. If this talk should be accepted, whichever option sounds most pleasing to the organizers is fine with me.

2 My own motivations and background

I was contacted by some of the workshop organizers for a survey about my maintainership of mlpack over the years. After a long and enlightening discussion, they encouraged me to submit an abstract to the workshop.

Personally, I would be interested in participating in some type of hackathon event; however, I am not convinced at the moment that the difficult barrier to entry of a library like mlpack can be solved in such a way that hackathon participants can quickly go from zero to contribution in a few short hours with little background knowledge. As I am not familiar with the research community studying hackathons, I am interested to see what types of insights and lessons I can learn, in order that I can either move forward and hold a hackathon-type event for mlpack, or even bring those insights and lessons back to mlpack's participation in Google Summer of Code and related programs.

3 Some minor background on mlpack

The mlpack machine learning library is a C++ project that has been in active development for nearly a decade. Originally, the project was a collection of code from a research lab at Georgia Tech; however, this was refined and released in 2011 as a collection of general-purpose algorithms that could be used by researchers anywhere. In the time since its release, mlpack has been cited in numerous academic papers and downloaded over 50000 times. mlpack has participated in the Google Summer of Code program in 2013, 2014, and 2016 (it was not accepted in 2015), and has mentored a total of 14 students in that time.

The mlpack website can be found at `http://www.mlpack.org`, and project development proceeds on Github at `https://github.com/mlpack/mlpack`.

CSCW 2017 Hackathon Workshop Submission: Collegiate Hackathons as Liminal Spaces

Erin R. Hoffman
2nd-Year PhD Student
Human Centered Design & Engineering
University of Washington
Seattle, CA 98105, USA
erinrhof@uw.edu

Abstract

Collegiate hackathons – time-bounded innovation competitions that take place on college campuses and are organized and attended by college students – are a large and growing phenomenon across the United States and the world. In the past year alone, thousands of student organizers have raised millions of dollars from corporate sponsors in order to host tens of thousands of student attendees. When describing their motivations, these organizers frequently refer to the "hacks" (projects created at hackathons) or the educational experience for "hackers" (hackathon attendees) as the purpose of their hackathon. Based on a qualitative study of hackathons, I argue that hackathons are also a liminal space between computer science education and the technology industry, where hackers embark on the transition between identities as "students" and identities as "employees".

Author Keywords

Collegiate Hackathons; Computer Science Education; Grounded Theory

Introduction

I have been a collegiate hackathon attendee and organizer since 2014 and recently performed a

Hackathon Timeline (My First Hackathon)

12noon Friday: Board a bus in East Lansing, Michigan

4pm Friday: Arrive at Purdue University in West Lafayette, Indiana. Drive around campus twice. Eventually realize that Purdue hackers have been picked up by another bus.

8pm Friday: Finally arrive at HackIllinois.

10pm Friday: Opening ceremony, then start hacking! Work on project, go to tech talks, snack, play cards, sleep on the floor, go for a walk, etc.

10am Sunday: Stop hacking! Practice pitches, grab lunch, pitch project to judges.

4pm Sunday: Closing ceremony. Top 10 projects pitch to entire crowd.

10pm Sunday: Arrive back in East Lansing, having done no homework but at least catching a nap on the bus.

Grounded Theory investigation of hackathon organizing. I am thus both a hackathon practitioner and researcher. I hope to contribute to the 2017 CSCW "Hacking and Making at Time-Bounded Events" workshop in the contexts of *design variations*, *short-term and long-term outcomes*, *practical support for event organizers*, and *theoretical space of hackathons*.

Experience as a Collegiate Hackathon Attendee

I attended my first collegiate hackathon, HackIllinois, at the University of Illinois at Urbana-Champaign in April 2014. This was the final month of my junior year as an undergraduate Computer Science and Engineering (CSE) student at Michigan State University (MSU). HackIllinois was a 36-hour hackathon, meaning that my weekend involved boarding a bus around noon on Friday, "hacking" (working on building our hack) from 10pm Friday until 10am Sunday (sleeping nights on the floor of a brightly-lit classroom), demoing our hack on 4pm Sunday, and arriving back in East Lansing around 10pm Sunday night. See the side bar for a more detailed timeline.

This ridiculous-seeming event had taken up my entire weekend, but I wasn't upset about it. My team member and I had built a "real" app in just 36 hours, we had demoed in front of hundreds of people, and our shared trials and tribulations (spending so many hours on a bus, sleeping on the floor, and neglecting our homework for so long) meant that the group of 30 or so students from MSU had developed a real sense of camaraderie. We were having the time of our lives!

In the following year I attended seven additional collegiate hackathons, spanning the country from HackMIT in Massachusetts to TreeHacks at Stanford in California. My teammate and I began strategizing to build hacks that would win the most prizes; one highlight was when we won five different prizes (valued at around $1000 total) at MHacks IV at the University of Michigan. On two occasions we made it to the top three, meaning that we were awarded medals by MLH and gained special access to "celebrity" hackathon judges like notable computer scientist Jeff Dean. In the spring of 2015, we were offered a space on a reality show about hackathons being filmed at HackDFW in Dallas and being judged by celebrity investor Mark Cuban (snow at DFW airport unfortunately foiled this plan). In addition to silly prizes like drones, LEGO sets, and swords, the connection requests and job opportunities from these hackathons' corporate sponsors started pouring in. Though I was awaiting graduate school admissions decisions and therefore off the job market, my teammate was able to take advantage of these opportunities. Today she works at Google and unambiguously attributes her career trajectory to the opportunities provided by hackathons.

Experience as a Collegiate Hackathon Organizer

Soon after our first trip to HackIllinois, a number of MSU students began considering how we might found a hackathon at our school. I joined as Facilities Director, and my hackathon teammate took the lead as Executive Director. From the start, SpartaHack's success looked unlikely. When we walked into the Student Services office, we were informed that the university did not permit overnight events, period. The coordinator of our CSE capstone program went out of his way to prevent its corporate sponsors from becoming our corporate sponsors. Our Sponsorship Director dropped out of school and moved to China, and our Webmaster attempted a coup to replace our

What Defines a Collegiate Hackathon?

Major League Hacking (MLH) sets these standards [5] but many non-MLH events also follow them.

Only for students: Often only includes college undergraduates. Sometimes includes high school (or even middle school) students or graduate students. "Student" label lasts for 12 months after graduation.

Free to attend: MLH Member Events may not charge an admission fee, and must provide all food and drinks. Funding comes from corporate sponsors.

24 to 36 hours: This is just the hacking time itself. With opening and closing ceremonies, these events take all weekend.

Prizes awarded: Hacks can win prizes for things such as "best Microsoft hack," "best first time hack," or "2nd place overall." Prizes are awarded by sponsors or by the hackathon itself.

Executive Director when she refused to book an expensive, too-small off-campus venue.

Eventually, however, things started coming together. Based on the financial endorsement of our first few corporate sponsors and the sanctioning endorsement of the student hackathon governing body Major League Hacking (MLH), the university finally permitted us to book the building with the campus's largest lecture hall overnight for the duration of the weekend. Members of our organizing team paid thousands of dollars with personal credit cards for t-shirts, buses, and catering. With the help of a new Budget Chair, we raised the majority of the budget for the event in the last month leading up to it. We were brimming with pride and relief when SpartaHack opened in March 2015 with 300 attendees, 18 corporate sponsors, a $46,000 operating budget, and $23,000 in prizes.

The following year, I began my PhD program in Human Centered Design & Engineering at the University of Washington (UW). In my first year at UW I volunteered with their student hackathon, DubHacks, then in the second year I officially joined the DubHacks team as a Logistics Director. We attended HackCon, MLH's hackathon organizer conference, as a team in the summer of 2016 and hosted DubHacks itself in October 2016. While I was a member of this team, I took a graduate course in Grounded Theory and applied these methods to observations of my hackathon organizing team. The following research abstract is the result of these observations.

Collegiate Hackathons as Liminal Spaces

Grounded Theory is a qualitative research method and is generally considered to be a form of ethnography [4]. Part of the process of Grounded Theory is sharing data (such as notes taken by an observer) with a research team and collectively identifying "codes" that appear repeatedly. The following analysis stems from one of the codes my classmates and I identified, which at the time we named "entitlement clashing with economic realities."

Broadly, I argue that collegiate hackathon organizers present the purpose of hackathons primarily as the hacks themselves, and secondarily as the educational opportunity for hackers. I believe that a third component, employment, is largely left unmentioned by hackathon organizers because acknowledgement of capitalism in a university setting clashes with "free culture" norms associated with hackathons and the tech industry more broadly.

The Purpose of Collegiate Hackathons as Described by Collegiate Hackathon Organizers

The MLH frequently asked questions (FAQ) page offers this response, which focuses almost entirely on hacker education, to the question of "What is a hackathon?":

"A hackathon is best described as an 'invention marathon'. Anyone who has an interest in technology attends a hackathon to learn, build & share their creations over the course of a weekend in a relaxed and welcoming atmosphere. You don't have to be a programmer and you certainly don't have to be majoring in Computer Science" [5].

In other cases, hackathon organizers focus more on the hacks themselves, either as forms of innovation or as hacks "for good" working toward a charitable cause. The DubHacks website, for example, states that "We

believe technology has the power to spark positive social change in our communities" [3].

The Purpose of Collegiate Hackathons as I Observed

Given my sensitizing ethnographic course work and observation of the large budgets involved in these events, I argue that they are primarily serving as liminal spaces between the world of the student and the world of the employee. In the structured environment of a collegiate hackathon, students interact with real-world technologies and representatives of real-world employers in a way that would be frowned upon as the "educational industrial complex" [2] if it were offered directly by universities. This explains why corporate sponsors are willing to spend so much money; they are buying their way into a formative experience for students. And 85% of hackers, when asked, report that they are hoping to get a job and/or internship by applying to this year's SpartaHack [Hammerly, private communication].

This economic reality does not sit well with many hackathon organizers, many of whom express distaste for corporate involvement. There seems to be an implication that the need for corporate sponsorship corrupts the otherwise pure hackathon experience.

I see this rejection of the capitalist nature of collegiate hackathons as having two causes. First, I believe that this is linked to the larger "free culture" movement, which advocates for things like free and open source software and also promotes ad blockers and bemoans the presence of advertisements on sites like Google and Facebook. This movement fundamentally believes that certain things, like internet access and hackathons, should be free with no strings attached. I believe that this belief is closely tied with arguments about meritocracy, since those who gained success through things like hackathons would feel less "self-made" if hackathons weren't free and open. Second, I believe that this is linked to technosolutionism, the idea that technology can and should solve the overwhelming majority of our problems. I see this in examples like the recent HackPrinceton project that claimed to have solved the problem of fake news online in just 36 hours [1]. If hackathon organizers subscribe to technosolutionism, it is understandable that they would be willing to focus on the hacks and disregard the liminal economic context.

References

1. Julie Bort. 2016. It only took 36 hours for these students to solve Facebook's fake-news problem. *Business Insider*. Retrieved December 18, 2016 from http://www.businessinsider.com/students-solve-facebooks-fake-news-problem-in-36-hours-2016-11
2. Dipti Desai. 2015. Educational Industrial Complex. Retrieved December 18, 2016 from https://greatschoolwars.files.wordpress.com/2015/10/eic-oct_11.pdf
3. DubHacks. FAQ. Retrieved December 18, 2016 from http://16f.dubhacks.co/
4. Kathy Charmaz. 2014. *Constructing Grounded Theory*. Sage.
5. Major League Hacking. FAQ. Retrieved December 18, 2016 from https://mlh.io/faq

Hacking at Microsoft

Hackers collaborate on a project at the 2015 global Hackathon in Redmond.

This paper is written by The Garage Hackathon team (www.microsoft.com/garage) to provide observations of the Microsoft Global Hackathon and hacking activities at Microsoft.

Hacking at Microsoft: An Introduction

Microsoft employees have been hacking individually and as teams for several years. In 2014, as part of updating and increasing employee engagement, a global, company-wide hackathon was developed. The Microsoft Global Hackathon is a highly visible demonstration of Microsoft's commitment to employees in support of the values of a hack culture. In its first three years, the Hackathon has received exceptional satisfaction and value ratings from its participants. Now the largest private corporate hackathon on the planet, the event is still, at its roots, an opportunity for employees to hack on something that interests and inspires them. The Global Hackathon is going into its fourth year, and many other hacking activities have also grown in popularity across the company as a "hacking culture" becomes integrated into the way Microsoft employees approach their day-to-day work. Year-round hacking activities exist to cultivate employee innovation, inspiration, and collaboration.

Participation is voluntary, and there have been several Microsoft internal tools and methods developed to enhance team-building and skill-building. Special attention is given to designing hacking activities and participation that is available to employees and interns across the company and across the globe.

Microsoft Hacking Fast Facts

Over three years:

- Registered 42,035 participants for hackathons (31,895 unique Microsoft employees)
- Registered 3020 college interns
- Registered 10,645 projects
- Registered hackers in 139 different cities and 75 different countries

- Developed an internal website/tool to support hacking, hackers, and hack projects. The internal tool supports creating hackathons, hacker profiles, and hack projects as well as resources to assist hackers find projects and projects find hackers via a skills, roles, and interests matching. There is even a bot to automate the process. The tool also provides technical and non-technical resources to support hackers and hack teams at every stage of the process. Online voting and online participation is tracked for each project.

Tools & Partnerships
An online internal tool called "HackBox" was created to support the Global Hackathon. It soon became clear that there was a demand for the tool's functionality to support year-round hacking through project creation, team building, skills matchmaking, and other collaboration. In 2016, an intern team added "HackBot," a bot that helps match hackers to projects and projects to hackers. It proved to be a great way to increase the site's impact.

The visibility of the Global Hackathon has brought about many partnerships across the company, including sharing of developer resources and non-technical resources outside of typically siloed processes. Subject matter experts want to share their knowledge and likewise, hackers want to learn and share their ideas. Functional teams have also rallied to help the Hackathon events team with resources like demo video uploading and playback.

Desired Outcomes
All ideas are welcome at the Microsoft hackathons, with a growing number of examples of how hack projects have been directly and indirectly delivered to customers. There is also an emphasis on the hacking exercise to evolve the company to a more 'growth mindset' culture. And participating in the Hackathon means learning from teambuilding, collaboration, meeting new people across the company, picking up new skills, getting inspired by colleagues — all less measurable but equally important outcomes.

Recognition, Judging, Winning
Competition is an element of the Global Hackathon, although it is not a priority of the event organizers. There are hundreds of recognized project winners, including category winners, executive challenge winners, sponsored challenge winners, online voting winners and others. Judges include both subject matter experts and peers.

Hacking as a means to move ideas forward
Hackathons are usually time-bounded events, but the hacking culture is becoming more and more integrated in the way employees work. Hackers are empowered to continue to work on their projects and ideas, and to seek the best ways to move their projects forward. There is an internal channel for releasing certain projects to customers (see examples at www.microsoft.com/garage)

Understanding Hackers at Microsoft

There are two generalities about Microsoft hackers that stand out. First, the number of people who hack is notable: there have been nearly 32,000 employees and interns who have participated since the first Hackathon in 2014. Second, the diversity of hacker profiles continues to be impressive: hacking continues to attract people from across multiple regions, roles, and groups at Microsoft. Through deliberate program and outreach design, the hacking opportunities and formats appeal to a broad range of participants and continue to receive high satisfaction ratings from all types of hackers.
For example, there are hackers from different --

- Professions and disciplines: engineers, developers, designers, researchers, program managers, marketers, sales, finance, evangelism, human resources, and legal and corporate affairs.
- Tenures: ranging from 1 week to 31 years at Microsoft
- Global locations: there are participants from more than 100 countries, and they hack with teams that are local as well as global. At the 2016 Microsoft Global Hackathon, there were 11 official venues in the US and another 32 worldwide.

One other noteworthy point about hackers besides their numbers and diversity: over the years, there are more examples of hackers with advanced hack skills. In addition to project work, advanced hackers assess market need, research customer and market value, execute work items that support customer deployments (e.g., includes compliance, legal review, supportability), and proactively build hack teams to include members with diverse skills and perspectives.

Understanding Hack Projects at Microsoft

At Microsoft, there is no shortage of ideas for hack projects. There have been more than 10,000 ideas generated for hack projects in just three years — with the number increasing each year — demonstrating that Microsoft employees welcome an outlet for pursuing personal ideas.

Projects are proposed for a variety of desired outcomes. Primary outcomes include impacting Microsoft customers directly with new products/services or improvements to existing products/services. Some projects are proposed to improve Microsoft employees work life (e.g., improve an internal tool, system, procedure, or process). Other projects are proposed as an opportunity learn a new skill, collaborate with peers and colleagues in different workgroups, or to spend hands-on time becoming familiar with a new technology. It's important to note that projects are not required to be technical. While most projects are indeed technical (i.e., involve code development), there are many examples of impactful non-technical projects (e.g., hack a new outreach program, business process, or internal policy).

Building diverse hacking project teams matters. Teams that include diverse skills and have members from different work groups are more likely to have a project that succeeds, win recognition, and have potential to ship to customers. Some projects are done by individuals or small pre-formed teams. Most project teams understand that collaboration and diverse teams contribute to successful outcomes, and proactively build teams with employees outside their immediate workgroup and with diverse roles and skills.

A popular project theme at Microsoft is 'HackForGood', where projects are specifically targeted at to demonstrate how technologies and Microsoft employee skills can help solve societal problems. In the recent 2016 Microsoft Global Hackathon, non-profits partnered directly with hack teams to align impact to local communities.

There are now more examples of advanced hack projects. Advanced projects target customer and business growth, have completed customer trials, have integrated work for market-ready deployment (e.g., compliance, legal review, supportability), and have expanded team members to include business and go-to-market roles.

Why Do Microsoft Employees Love Hacking?

Hacking is **Microsoft affirming.** Hackers embrace the time to explore and learn technologies that are aligned to Microsoft strategies. Building projects, and seeing other project demos built with these

technologies provides tangible examples of exciting capabilities and helps employees get excited about Microsoft investments and strategies. Additionally, projects that exemplify how Microsoft technologies and employee skills can contribute to societal good re-affirms Microsoft's long-standing commitment to philanthropic efforts.

Hacking is **idea affirming**. Hacking at Microsoft provides an opportunity for anyone to offer a project idea, more importantly, act on it. That anyone, regardless of level, role, and tenure, can create a project and build a team, reinforces that all ideas are welcomed. That Microsoft has a mechanism that represents a true democratization of innovation amongst employees is distinct in corporate culture and provides a motivating proof point that 'having Ideas' is alive and well at Microsoft.

Hacking is **employee affirming**. The sentiment among hackers is that hacking experience makes great use of their total skillsets. In just a few days, hackers can get hands-on experience in all stages of project development and value the opportunity to flex skills in areas such as research, testing, feedback, and marketing. The ability to wear many hats contributes to the perception of broader personal skills and the appreciation of the diverse skills of teammates. For many, the Hackathon becomes the first exposure to diverse skills and roles at the company. The most common reaction to viewing hack projects is to be inspired by the energy and creativity of colleagues. Overall, the hacking affirms that Microsoft employees are passionate, creative, and innovative.

Hacking is **action-orientation affirming.** Hacking is fundamentally about "doing." Hackers create the minimally viable structure to get team members working and spend most of their time hands-on with projects and teammates. It's personally gratifying to make substantial progress in a short amount of time. Producing a simple demo video that can viewed worldwide produces pride in hack teams and in a job completed.

Hacking is **personal empowerment affirming**. Participating in hacking provides opportunities for leadership that may not be available in day-job workgroups. Teams with diverse skills, roles, and employee levels collaborate equally and without hierarchy. But everyone on the hack team also can be a manager, innovator, and leader. Hackers feel empowered by their own ideas, work from their own enthusiasm, and have considerable pride in making something that reflects personal passions and interests. In particularly, the freedom to create and control their hack work environment is considerably empowering.

Hacking is **personal development affirming**. Hacking lets people explore experiences and skills outside their day job. Working through all project phases (idea creation, project definition, implementation, refinement, messaging, marketing) provides hands-on opportunity to acquire and practice research, marketing, and business skills. Everyone gets to try out the role of maker, builder, planner, business person, marketer, and salesperson.

Hacking is **cultural change affirming**. Microsoft employees have heard messages about culture change from all levels of the company. For many, the hacking provides a tangible and visible example of Microsoft's commitment to evolving culture. The worldwide scale, dedicated time, resources made available, and cross-company participation provide highly visible proof points that back up the messages from leadership. Hack team experiences move employees from passive listeners to active contributors, where the project teams demonstrate core concepts of change (new ideas from all levels, experimentation, learning, building relationships outside workgroups, valuing different skills, and perspectives). Hacking moves employees from passively hearing about cultural change to actively participating in culture change.

Community and Code: Lessons from NESCent hackathons

Arlin Stoltzfus

I am a full-time professional researcher. After receiving laboratory training in molecular and population genetics, I turned to computer-based studies of molecular evolution. I develop and explore computer-based models, test hypotheses using available data, and collaborate with others to develop software and standards that improve the interoperability of evolution-related data and services.

From 2006 to 2015, I was involved with a series of 9 hackathons sponsored by the National Evolutionary Synthesis Center (NESCent), an NSF center dedicating to promoting integrative and synthetic work. My involvement was as a participant, as an instigator (securing support), as an organizer (planning and recruiting), and as a facilitator (managing the hackathon process). On several occasions I organized or participated in post-hackathon campaigns to produce a publication or a grant proposal.

NESCent hackathons are distinctive in format, aims, and audience. The community of computational evolutionary biologists is a small, geographically dispersed, poorly resourced group. NESCent hackathons were destination

events (and not, for instance, conference workshops) that subsidized the cost of participation (each hackathon had a budget of $25K to $30K spent largely on travel). Each event had a theme defined in terms of technology, domain-specific challenges, or both. At each event, 25 to 35 participants met on the first day to participate in a staged process culminating in facilitated bottom-up team-formation, then worked for another 3 or 4 full days in teams to produce tangible outcomes. NESCent's aims in sponsoring these hackathons was community-oriented, focused both on nurturing a community of practice in evolutionary informatics, and on generating broadly useful products.

Because NESCent hackathons generated considerable enthusiasm and seemed to be highly successful, a group of individuals involved in the events undertook to document the NESCent hackathon model, and to conduct a retrospective analysis. Because the hackathons were organized as "Open Science" events, with open-source code repositories and team documents on public wikis, a vast amount of information remains publicly available on the internet. From this information, we compiled a description of 9 hackathon events, 54 teams, and 148 products (mostly team reports and code repositories). This information, along with our personal recollections, served as the basis for reflections on hackathon processes, outcomes, and impacts.

The most obvious tangible outcome of NESCent hackathons was computer code, often consisting of stand-alone proof-of-concept software, but also frequently

involving incremental additions to pre-existing code-bases. Less often, hackathon teams produced documentation, designs, installations, or data products. Typically, when the hackathon ends, and team members disperse and return to their day jobs, the team's code repository and other tangible outcomes become inactive: no further work is done and there is apparently no direct productive use of the hackathon products.

In terms of tangible outcomes and their impacts, the value of hackathons lies in the exceptions to this general rule. During the wrap-up to one hackathon event, participants decided to create a mailing list that, 10 years later, has over 1000 subscribers and 30 to 60 messages per month. In a minority of cases, post-hackathon efforts lead to downstream outcomes such as blogs, presentations, manuscripts, and proposals for funding. NESCent hackathons have so far led rather directly to 4 publications and to 2 grant proposals securing over $2M of funding to develop systems based on proof-of-concept software. The case for a positive return-on-investment for hackathons could be made from these two grant proposals alone.

Furthermore, the direct impact of tangible outcomes does not tell the entire story. First, tangible outcomes may have intangible impacts. For instance, even if a proof-of-concept developed at a hackathon is unused and is not the proximate basis for any further work, the developer may go on to implement an improved or re-designed version, benefitting from the experience gained by the first

implementation. Second, the hackathon has many intangible outcomes. For participants, intangible outcomes include learning new technologies, exposure to best practices, team programming, and increased awareness of shared challenges and opportunities facing the community. Hackathons often provided a venue for participants to take risks, foregoing ordinary demands of productivity to experiment with a new approach that might not pay off.

Our experience suggests several lessons for organizers, at least for the case of hackathons that share some of the aims of NESCent hackathons. The declared theme and scope of the hackathon should be defined clearly and communicated unambiguously, and must combine the potential to inspire with flexibility in interpretation. Providing opportunities for pre-event engagement with participants is important, even if participation is partial. An effective strategy for increasing diversity is to identify qualified candidates from under-represented groups and encourage them personally to apply (the rate of return on personal appeals is very high). To facilitate development of an open record of freely shared information, it is important to establish in advance a set of preferred technologies for code management, document-sharing, and communication.

This experience also raises questions. As a hackathon organizer, my main interest is in organizing events (scoping, recruiting, event logistics, followup) so as to increase hackathon effectiveness and maximize return-on-investment for the sponsor. My main interest as a hackathon facilitator is to maximize

participation (and particularly to ensure that participants who are not highly networked are included and not left behind), and to ensure that poorly conceived teams (projects) do not emerge. However, in regard to effectiveness, it is not clear what this means. The obvious tangible outcome of a hackathon is computer code, but our experience bears out the frequent complaint in the literature that most hackathon code is a dead end. To the extent that hackathons are primarily networking events, we should be assessing their effectiveness relative to social gatherings such as parties. To the extent that they are learning events, we should measure their effectiveness against workshops or lectures. However, it may be that hackathons are a unique combination (learning-and-doing parties?) that cannot be compared to anything.