Check for updates

RESEARCH ARTICLE

Community and Code: Nine Lessons from Nine NESCent Hackathons

[version 1; peer review: 1 approved, 1 approved with reservations]

Arlin Stoltzfus[1], Michael Rosenberg[2], Hilmar Lapp[3], Aidan Budd[4,5], Karen Cranston[6], Enrico Pontelli[7], Shann Oliver[8], Rutger A. Vos[9]

[1]Institute for Bioscience and Biotechnology Research, University of Maryland, 9600 Gudelsky Drive, Rockville, MD, 20850, USA
[2]School of Life Sciences, Arizona State University, Tempe, Arizona, USA
[3]Center for Genomic and Computational Biology (GCB), Duke University, Durham, NC, 27708, USA
[4]Structural and Computational Biology Unit, European Molecular Biology Laboratory (EMBL), Meyerhofstraße 1, 69117 Heidelberg, Germany
[5]Earlham Institute, Norwich Research Park, Norwich, NR4 7UG, UK
[6]Biology Department, Duke University, Durham, NC, 27708, USA
[7]Computer Science Department, New Mexico State University, Las Cruces, NM, USA
[8]Guardant Health, Redwood City, CA, USA
[9]Naturalis Biodiversity Center, Vondellaan 55, 2332 AA, Leiden, The Netherlands

v1 **First published:** 06 Jun 2017, **6**:786
https://doi.org/10.12688/f1000research.11429.1
Latest published: 06 Jun 2017, **6**:786
https://doi.org/10.12688/f1000research.11429.1

Abstract

In recent years, there has been an explosion in the popularity of hackathons — creative, participant-driven meetings at which software developers gather for an intensive bout of programming, often organized in teams. Hackathons have tangible and intangible outcomes, such as code, excitement, learning, networking, and so on, whose relative merits are unclear. For example, a frequent complaint is that code is abandoned when the hackathon ends, and questions like, “which outcomes are produced most reliably?” and, “how valuable are they for participants, organizers, and sponsors?” remain open. As a first step in giving “hackology” a more rigorous footing, this paper describes the NESCent hackathon model, developed over the course of a decade to serve the academic discipline of evolutionary biology, with the dual goals of augmenting the community’s shared software infrastructure, and fostering a diverse community of practice in open scientific software development. The paper presents a detailed guide to staging a NESCent-style hackathon, along with a structured information set on nine events involving 54 team projects. NESCent hackathons have produced tangible products with downstream impacts, including prototypes that were leveraged for major funding, incremental additions to production code bases, and creative drafts (designs, standards, and proofs-of-concept) that contributed to further work. The impacts of intangible outcomes could not be assessed objectively, but the subjective experience suggests that hackathons have a positive impact by (1) providing individuals with

Open Peer Review

Approval Status ✓ ?

	1	2
version 1 06 Jun 2017	✓ view	? view

1. **Eva Amsen**, The Institute of Cancer Research, London, UK

2. **Cameron Neylon**, Curtin University, Perth, Australia

Any reports and responses or comments on the article can be found at the end of the article.

valuable experiences, networking, and training, and (2) fostering a more cohesive community of practice by enhancing awareness of challenges and best practices and by building links of familiarity between and among resources and people. Future research that recognizes the diverse outcomes of hackathons might enable evidence-based decisions about how to design hackathons for effectiveness.

Keywords
hackathon, programming, software development, scientific software, NESCent



This article is included in the Hackathons collection.



This article is included in the Bioinformatics Education and Training Collection collection.

Corresponding author: Rutger A. Vos (rutger.vos@naturalis.nl)

Competing interests: No competing interests were disclosed.

Grant information: The National Evolutionary Synthesis Center (NSF #EF-0905606) extended its support for the HIP (Hackathons, Interoperability, Phylogenies) working group by funding the authors to gather at NESCent in March of 2015 to lay the groundwork for this project.
The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2017 Stoltzfus A *et al*. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Stoltzfus A, Rosenberg M, Lapp H *et al.* **Community and Code: Nine Lessons from Nine NESCent Hackathons [version 1; peer review: 1 approved, 1 approved with reservations]** F1000Research 2017, **6**:786 https://doi.org/10.12688/f1000research.11429.1

First published: 06 Jun 2017, **6**:786 https://doi.org/10.12688/f1000research.11429.1

Introduction

Hackathons (also called *hackfests* or *codefests*) are short-term software development events that emphasize spontaneity and collaboration, bringing together developers, and sometimes end-users, with the goal of innovative software development, often in conjunction with other objectives such as fostering a community (i.e., building a stronger "community-sense"[1]), or drawing attention to particular data or services. Since the early 2000s, hackathons have become increasingly popular (Figure 1) - including across academic, non-profit, corporate, and government sectors - with events focused on a variety of topics, such as bioinformatics[2], promoting open data[3], medical education[4], and healthcare informatics[5].

Although a lot of information on hackathons can be found online - including various guides[6,7] and reports on specific events - there is very little academic, peer-reviewed literature on the topic. Of the small amount of published work available, most consists of reports on specific hackathon events, some of which are short[4], while others go into depth about the technical products of the event rather than the process[2,8–11]. Others are news and opinion pieces[12–15]. Only a few published sources are based on systematic methodology such as surveys, interviews, or organizing structured data[16–20].

Thus, in spite of the popularity of hackathons, there is currently no systematic basis for evidence-based approaches to planning or organizing a hackathon. For a prospective organizer, the immediate practical question is how best to carry out a hackathon. If we assume that a hackathon is typically carried out with the intent to maximize its benefits to its sponsors and its participants, then the question of how to conduct a hackathon requires understanding these benefits, and more generally, understanding why hackathons are carried out at all.

It turns out that there is no clear consensus on exactly how hackathons bring value to participants or sponsors. For example, although the obvious expected outcome of a software development hackathon is software, organizers frequently note that these events generate prototypes, not products used after the event[21,22].

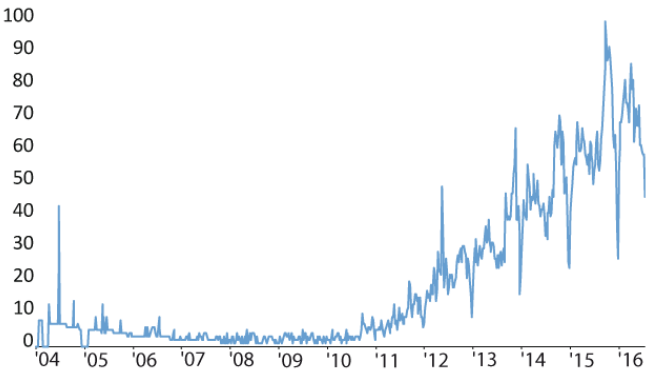


Figure 1. Weekly Google Trends results for the term "hackathon" from January 2004 through to June 2016. Values are relative to the highest point on the chart, thus the week with the greatest search interest in the term receives 100% and other weeks are scaled accordingly.

If the source code generated at hackathons is rarely used, then why are hackathons so popular? One possibility is that, even if only a small fraction of code remains useful, this small fraction may still justify the event. Another possibility is that the benefit of hackathons arises partly or largely from less tangible outcomes. When a hackathon is focused on utilizing a sponsor's newly released API, the event may uncover bugs, or bring valuable exposure to the sponsor's resources or products (e.g. as in some of the hackathons described in 22). Even if a prototype developed at a hackathon is never used, the developers may leave the event with the experience and confidence to build a similar (perhaps improved) implementation later. Participants may benefit from gaining technical skills, from sharing best practices, and from making connections with colleagues, i.e. professional networking. For example, participants of the BioHackathon series of events[23–26] are strongly encouraged by the organizers to connect with each other on social networks, such as LinkedIn.

Not only direct participants themselves, but also the community they belong to may benefit from discussions and interactions that spread technical knowledge and create a shared awareness of domain-specific challenges, opportunities, and best practices[27]. The expectation of stimulating creativity and building camaraderie seems to be one of the motivations of internal hackathons (e.g., 22). In addition, participating in a community event may promote "collaborative learning", which is one of the top two reasons for attending a hackathon, according to participants cited in a recent publication[28], the other reason being networking.

Arguably, how a hackathon event is organized and executed will affect how the beneficial outcomes of hackathons, tangible and intangible, are enhanced or diminished. Indeed, hackathons vary in many ways, even within the broad categories of corporate, community, and internal hackathons[13]. They may be one-off events[29], or a series that repeats yearly[23–26] or even more frequently[15]. The event may last a single day (e.g., 12), an entire week[9], or longer[18]. The number of participants may range from a few dozen (e.g., 8, 27) to hundreds. Some events offer prizes[14,30]. There is considerable variety in how development targets are determined (e.g., 5) and how teams are formed[19,31]. Some events are carefully planned for months[10], while others emerge more spontaneously.

Hackathon organizers frequently establish a process to engage participants in learning, socializing, or brainstorming prior to the event[10,29,32,33]. For most hackathons there are no planned follow-up activities, but in some contexts (e.g., internal hackathons), resources may be set aside to build on promising outcomes[15,34]. In light of the extensive variability of hackathons, better information - and ultimately, systematic studies - on hackathon practices, outcomes, and impacts will be needed to better understand how and why to conduct a hackathon. To begin laying the foundations for a more systematic understanding, we offer a description and analysis of a series of relatively well-documented hackathons sponsored by

the erstwhile *National Evolutionary Synthesis Center (NESCent)*, an academic research center in the USA funded by the US National Science Foundation (NSF). Over a 10-year period, NESCent sponsored nine hackathons focused on software development to improve interoperability of software and data in the domain of evolutionary biology (comparative analysis, phylogenetics, etc.) (Table 1). Each event was planned by a leadership team whose membership intersects with the set of authors of this work, that is, each team included at least one of us, and most included several of us. The events all followed a common model for process and format, including length (4–5 days) and size (roughly 30 participants). The hackathons were designed both to develop tangible products and to foster a community of practice[35,36].

In the remainder of this paper, we begin with a detailed guide to the NESCent hackathon model, including the organizational process, and the motivations behind chosen practices. Then we describe known outcomes and impacts of the nine NESCent hackathons held, and reflect on some of the lessons learned as organizers and participants. Though our results on outcomes tend to confirm the sense that hackathon teams rarely produce novel prototypes that go on to be used, they often make incremental additions of code and documentation to production codebases that remain in use. In the rare event that novel prototypes and designs do contribute importantly to future work, the impact can be disproportionately large. Several hackathon projects led to publications, and two led to funding that exceeded the total cost of the hackathon by two orders of magnitude. Regarding intangible outcomes, although we lack sufficient data to draw firm conclusions, participants in NESCent hackathons seem to value the coding experience; they will have gained experience in problem-solving and teamwork, acquired training in supportive technologies, and improved their knowledge of best practices and awareness of resources, and opportunities for personal networking. NESCent hackathons also seem to build community by building operational links between community resources, creating excitement and a common focus of attention, and fostering cohesion and awareness with regard to best practices and domain-specific challenges.

Methods

The hackathons we describe (Table 1) were sponsored mainly by NESCent. As a consequence of the sponsor's commitment to open science, a large amount of information on NESCent hackathons was public from the outset. Agendas, slide decks, and other documents were developed and shared on public wikis; event rosters were shared publicly; teams prepared reports using public wikis, and were expected to share code in public source-code repositories. Most of this information has remained accessible on the web subsequent to NESCent's closure in May of 2015. From these sources, we have gathered a systematic set of data on NESCent hackathons, including data on (1) nine events (name, dates, scope, location, etc); (2) 54 projects (titles, descriptions); (3) 148 products (mostly team reports and repositories); and (4) numbers of participants (207 in total).

Data collection and encoding

The vast majority of information on the nine hackathons (time, place, theme, participant roster) and their team projects (goals, repositories, team reports) is available from public resources (e.g., wikis, code repositories). We also contacted participants to fill in gaps in this knowledge. In passing, we note that the quality of the available documentary record on NESCent hackathons decreases as one goes back in time, even beyond what one expects from the decay of records over time. It appears that participants in later hackathons were simply more effective at documenting their work, and organizers became more experienced in recognizing and emphasizing what kind of information needed to be documented. For example, the wiki for the first NESCent hackathon (phylohack, see Table 1) contained a relatively large amount of detailed planning information prepared before the event, but few specifics about what happened at the event.

In some cases, the interpretation of this source material requires judgment and domain knowledge, e.g., when a hackathon team did not provide a succinct statement of purpose or goals, we constructed a statement from the materials available, drawing on our recollections and our domain knowledge. In some cases the records left by participants made it difficult to distinguish prospective plans

Table 1. The 9 NESCent hackathons.

Short name	Date	Participants	Title	Sponsors
phylohack	Dec-2006	30	Phyloinformatics Hackathon	NESCent
comparativer	Dec-2007	28	Hackathon on Comparative Methods in R	NESCent
vocamp	Mar-2009	25	Phyloinformatics VoCamp	NESCent, TDWG
dbinterop	Nov-2009	31	Database Interoperability Hackathon	NESCent
gmodtools	Nov-2010	32	GMOD Tools for Evolutionary Biology	NESCent
phylotastic 1	Jun-2012	23	Phylotastic: the Tree of Life, as you need it	NESCent, iPlant, BioSynC
phylotastic 2	Jan-2013	27	Phylotastic 2	NESCent, iPlant, BioSynC
treeforall	Sep-2014	43	Tree-for-all Hackathon	OpenTree, NESCent, Arbor
rpopgen	Mar-2015	30	Population Genetics in R Hackathon	NESCent

from actual accomplishments, and in these cases, we used our best judgment.

While the information on events, projects, and participants is comprehensive (in the sense of describing - albeit incompletely - all events, projects and participants), we could not create a comprehensive list of products. This is partly because some products emerge long after the event, but also because teams sometimes produce several distinct products, but do not document all of them. The products that were easiest to find were (1) a team's report or activity log, as these were nearly always linked to the main web page for the event, and (2) the main code repository for a team.

It is much harder to track follow-on products. Examples of follow-on products include participants giving a talk at a conference, posting a blog, publishing a paper, or submitting a grant proposal based on hackathon outcomes or activities. To better characterize outcomes, we explored at greater depth a randomly selected set of nine projects, one from each of the hackathons. For each one, we sifted through online information and conducted a preliminary assessment of outcomes and impact, and then contacted a member of the original team to review the assessment and obtain further information, before settling on a list of outcomes and impacts. The entire dataset, as well as other supplementary material, is available at https://nescent.github.io/community-and-code/

Executing a NESCent hackathon

In this manuscript, individuals referred to as *instigators* get funding; *sponsors* support the event with funds and other resources such as space, logistics and IT staff; *organizers* participate in planning and making arrangements; *facilitators* manage group activities during the hackathon; *trainers* provide training on Day 1; and *participants* participate in Day 1 activities and join a hackathon team. A single person may play multiple roles over the course of a hackathon.

NESCent hackathons were five days (in one case, four days) long. For participants, the event was the main focus of attention and activity for the duration of the event. For organizers, by contrast, the hackathon was the culmination of a process that began months earlier when one or more instigators solicited support from sponsors, and assembled a Leadership Team (LT) of organizers to carry forward the planning process, recruit participants, and make all arrangements for a successful, well-facilitated event with sufficient training opportunities.

Table 2 provides a typical timeline of steps in the process, which might be accelerated in other contexts. Figure 2 illustrates the flow of the planning process. We do not provide complete guidance for prospective organizers in this article, but we refer prospective organizers to the *Concise Guide*, succinct instructions for planning and executing a NESCent-style hackathon, at https://nescent.github.io/community-and-code/doc/ These guidelines represent our current recommendations based on practices that have evolved over a decade. It is supplemented with other materials such as sample advertisements and application forms, all available in the online repository mentioned above.

Funding. Hackathons began with instigators who secured support based on a vision for a successful hackathon. These instigators typically came from NESCent's informatics staff or from one of two NESCent "working groups" (i.e., periodically convening collaborations among in-house NESCent staff and extramural researchers). NESCent was the sole or lead sponsor for most of the nine events. Sometimes there were grant-funded projects critical to the success of the hackathon that offered support, with the understanding that project staff would be participants in the hackathon. For example, the Phylotastic hackathon, which received support from the NSF-funded iPlant project, was staged at iPlant's home institution, and included a number of iPlant staff. The typical

Table 2. Organizational timeline for a NESCent hackathon.

Week	Milestone	Description
-26	Hackathon initiative	Concrete initiative to define scope, goals and LT
-17	Final scope	LT finalizes scope, goals, budget and logistics of hackathon
-14	Recruitment strategy	LT finalizes its recruitment materials, targets and timeline
-11	First RFP	Disseminate the solicitation for applicants
-9	Second RFP	Re-issue the solicitation for applicants
-8	Application deadline	Deadline for applicants to apply
-6	Review deadline	Reviewers to finish all applicant reviews
-5	Applicant decisions	Finalize decision on whom to invite
-4	Finalized roster	Definitive roster is determined and invites sent
-3	Travel arrangements	Finalize travel and accommodation plans for all participants
-2	Pre-event engagement	Initiate pre-event engagement with participants
0	Hackathon event	The actual hackathon event takes place
3	Report	Develop a report to the sponsoring organization

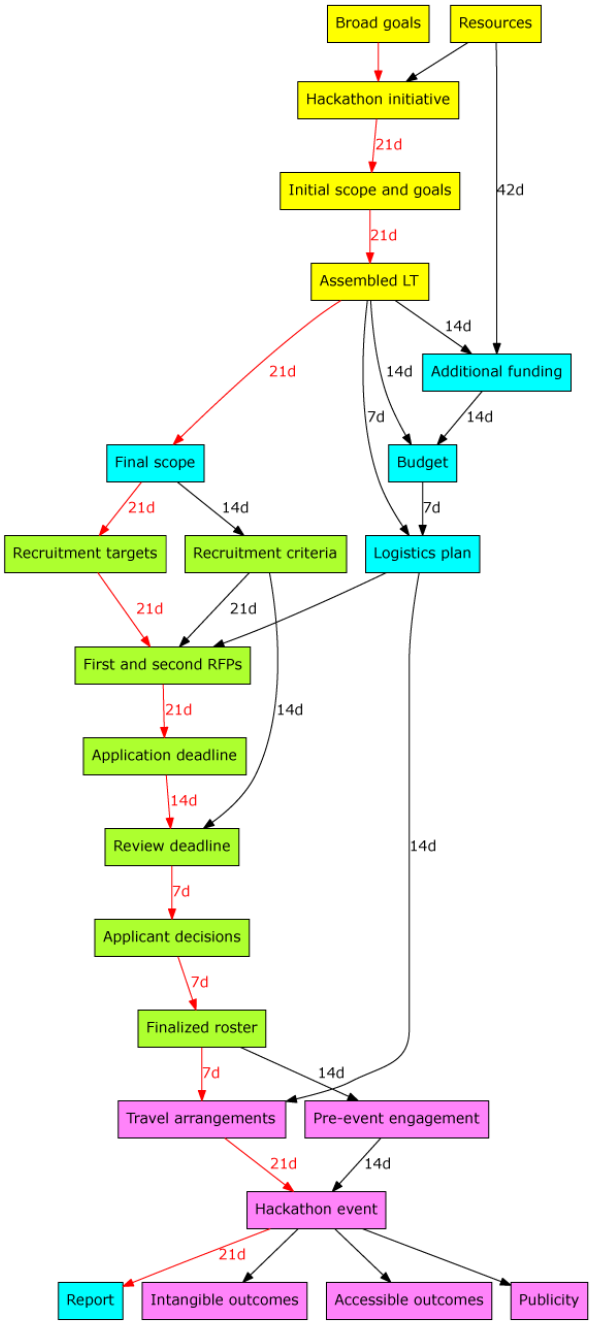


Figure 2. Program Evaluation and Review Technique (PERT) chart of the organizational process. The first steps (yellow) are taken by an informal group of instigators. Subsequently, a leadership team (LT) finalizes the pre-planning process (cyan), at which point the recruitment process starts. Inviting potential participants, reviewing and ranking their applications, and finalizing the roster are time-sensitive and labor-intensive steps (green), which lead up to steps that both LT and invitees participate in (fuchsia): the planning of the logistics and the actual substance of the event, including any follow-ups. As a final step, the LT reports back to any sponsoring organizations.

budget for a hackathon was $25,000 to $30,000 USD, nearly all of which was spent on travel. Meeting space was arranged with sponsor organizations at no cost. Sponsors also provided logistics support to arrange travel and on-site IT staff.

Planning. To initiate planning, the instigators recruited a leadership team (LT) of around five to seven organizers. To complete the entire planning process from scoping to finalizing a roster (i.e. the steps marked in green in Figure 2), the LT typically met five to eight times for hour-long teleconferences or videoconferences, over a period of two to three months. The motivation behind having a large team of organizers was partly to broaden decision-making, and partly to spread out the burden of making meeting arrangements, drafting advertisements, reviewing applications, and so on, in the absence of support staff.

LT members were chosen based on expertise, willingness to "think big", diversity, and expected effectiveness in hackathon planning. They were given an estimate of work hours expected (roughly one to two hours per week over the organizational period). Those who agreed to take part often had a keen interest in the topic of the hackathon and its potential to enhance their individual goals; for the team of recruited organizers (which are distinct from the instigators) to take true ownership of the project, they had to be allowed to re-think the scope and goals. Difficulty in assembling a committed LT, or in reaching closure on scope and goals, indicated a weakness in the instigators' vision (see "Lessons learned" section).

Next, the LT decided on a preferred set of supportive technologies for version control, shared documents, and communication. This made it easier for teams to collaborate, and for the LT to track progress and make sure all hackathon products were readily accessible. These choices changed over the years with changes in technology, e.g., early hackathons used SourceForge or Google Code repositories, while more recent ones used GitHub. We have used many technologies for creating and editing shared documents, including MediaWikis, Google Docs, Mozilla's Etherpad, GitHub documents, and others. In some cases, the use of a consistent document strategy resulted in a rich online record with links to code, screencasts, live demos, slides, etc.

The choice of communication strategies was most important before and after the hackathon. Email lists were an effective choice for organizers to convey plans, and also provided a forum for discussions in the pre-event engagement stage. NESCent organizers created two email lists that were used for multiple hackathons, with new participants added and prior participants retained. The choice of communication technologies used was also important to consider during a hackathon when remote participants were to be supported (see *Concise Guide*, at https://nescent.github.io/community-and-code/doc/).

Recruiting participants. Participants were either chosen from a pool of applicants responding to an open call for participation, or chosen directly and offered a seat at the hackathon. Dissemination

of the open call (and any advertisements) was done via email lists and websites that reached the target community, as well as by spreading word through emails to colleagues. Sample advertisements are included in Supplementary Material. The open call was a way to reach out broadly and engage unexpected members of the targeted community.

Over time, we relied less and less on direct invitations. In the most recent hackathons, we did not offer seats directly to anyone other than those organizers who wished to participate (organizers sometimes declined to participate so that someone else would be able to attend). Instead, individuals targeted for participation (for technical or diversity reasons) were invited personally to apply. This made the process more democratic, at the cost of occasionally not choosing someone who was invited to apply.

The application process typically was simple, though review of applications was the most time-consuming step in the organizing process. Over the years we developed a simple application form implemented as a Google form, allowing online entry of information that goes into a spreadsheet. Sample applications are available at https://nescent.github.io/community-and-code/doc/sample_applications, along with a link to a template that can be used to create an online form. Beyond basic information, we did not ask for much (see the *Concise Guide* at https://nescent.github.io/community-and-code/doc/ for an exact list). Two key parts were a statement describing qualifications (ideally with references to tangible accomplishments), and a statement of goals or aspirations for the hackathon.

Applicants for open seats were ranked according to estimates of expected impact on the success of the hackathon, taking into account that success of the hackathon requires teamwork, and may benefit from homogeneity in some areas (e.g. having a critical mass of people working on a particular topic) and heterogeneity in others (e.g. mixing users and programmers together).

Facilitating. Typically, two or three of the organizers served as meeting facilitators to guide participants through the hackathon process. In the weeks prior to the hackathon, we engaged participants with the aim of raising their comfort level, by introducing supportive technologies, providing a forum for discussion of ideas, and identifying gaps in technical or scientific knowledge. Our strategy and level of effort varied greatly; for example, at the repeat hackathon Phylotastic 2, less lead-up to supportive technologies or discussion of ideas was needed because these had already been established at the first one. In contrast, at the Database Interoperability Hackathon, novel technologies were introduced such as RDF. With repeated efforts and considerable prodding, organizers could get nearly everyone to join a mailing list or teleconference and introduce themselves to each other. In the more recent hackathons, we encouraged discussion via a GitHub issue list, which required the participants to sign up for a GitHub account if they did not have one already.

As many attendees were new to participant-driven meetings, facilitators repeatedly stressed that the event belongs to the participants and the teams they form; that each participant belongs at the meeting, and has a responsibility to become engaged so as to become part of a team where they are either contributing or learning.

The first day of NESCent hackathons consisted entirely of structured activities (see Figure 3). After a welcome and introductions, the organizers arranged for technical presentations on topics chosen based on the scope of the hackathon and the results of pre-event engagement. For instance, the TreeForAll hackathon focused on leveraging OpenTree's new web-services API, thus the organizers arranged for OpenTree staff to describe the API in the opening session of Day 1.

After these presentations, participants were engaged in an open discussion of ideas and challenges, with the aim of identifying a sufficient number of project ideas that were feasible and that aligned with the scope. Then the facilitators invited brief "pitches", project ideas proposed for broader adoption. Most pitches were anticipated based on earlier discussions. In practice, they often came from more senior people (including organizers) with a more confident sense of what projects would have an impact.

The champion for each pitch then created an impromptu poster. Participants were free to wander around the room, discussing pitches, offering suggestions, and deciding how to fit in. At this stage, the potential fit of a participant to a project is not like the fit of a key to a pre-existing lock, because the definition of the project is still in flux. Except in one instance in which the process carried over to the next day, the first day ended with a set of five to seven hackathon teams committed to a project with recorded goals.

The guidelines provided at https://nescent.github.io/community-and-code/doc/ outline the space requirements and room configuration for this team-development process. Some space configurations are inadvisable. A room with fixed stadium seating, for instance, is



Figure 3. Typical schedule for day 1 of a hackathon.

unsuitable, no matter how large. Other room configurations tend to create or amplify inequalities, e.g. a room with a single table large enough for most, but not all participants, will leave some without a seat at the table. A configuration where all team members can sit at the same table or tables such that they can interact easily without too much cross-talk from other teams works well.

The remaining days of the hackathon were spent with teams working on their individual projects, with pre-determined times for plenary sessions to hear team reports or "stand-ups". The stand-ups were meant to be short, and generally only needed to happen once a day, to avoid wasting too much time on updates. On the final day, stand-ups were skipped in lieu of a final plenary session to wrap up the meeting, typically including final team reports, along with discussion of possible products that might be achieved with minimal effort after the hackathon (e.g. publications, presentations at conferences, commits to codebases). Some wrap-up sessions included more general discussions about long-term follow ups (e.g. identification of potential funding sources that would enable scaling up of some of the development efforts).

Organizers typically carried out follow-up activities after the hackathon. They ensured that travel reimbursements were made, and they produced a report on the hackathon, which ranged from a blog to a manuscript for publication. Generally, very little could be expected of participants once they left the hackathon and went back to their "day jobs". However, the organizers sometimes interacted with participants to follow up on projects, e.g. to make sure that a team's report was going to be made available on a public web site.

Results and discussion

The results of a hackathon (Table 3) can be separated by outcome. For present purposes, we define outcomes as direct results of the activities of hackathon participants, whereas impacts are defined by how these outcomes penetrated in the larger world. We distinguish outcomes of the hackathon itself from outcomes of follow-on activities by participants. Also, outcomes may be tangible or intangible. For instance, code is a tangible outcome of a hackathon that can be counted (e.g. as lines of code, number of functions or objects), and the impact of the code can be assessed in terms of the number of times the code is invoked in a production setting or mentioned in online discussions. Typically, these outcomes result from the efforts of a specific hackathon team, but some outcomes result from the event as a whole. Thus we distinguish below between project (team) products (PP) and event products (EP).

As mentioned in the methods section, we took a closer look at nine projects (one at random from each hackathon). The remarks below illustrate what it looks like for hackathon products to have an impact. Additional cases have been added, whenever we happen to have knowledge of their outcomes and impacts. Of course, in those cases, our information on impact will not be systematic.

Tangible immediate outcomes and their impacts

Tangible hackathon outcomes included, in rough order of decreasing frequency: (1) new code repositories and incremental additions to existing code; (2) documentation; (3) designs, standards and schemata; (4) installations; (5) data products; and (6) community infrastructure.

Typically, the main product of a hackathon team is computer code. Hackathon code often represents a new project in a new repository. For instance, the last 3 events (the two Phylotastic hackathons and the Open-Tree hackathon) produced 15 new Bitbucket or GitHub repositories (PP#1 through PP#14, PP#17). The "Integrating Ontologies" group (project 18) from the VoCamp produced several different standalone products (PP#18 to PP#22), most representing an integration of an ontology with data, or with another ontology. Sometimes, developers familiar with an existing software package make additions or improvements that become part of production code. For instance, the first hackathon targeted improvements in existing toolboxes, including BioPerl, BioRuby, BioPython and Bio::Phylo (a Perl package separate from BioPerl, Vos *et al.*[37]). Other examples would be additions to Phylomatic (PP#119), Forester (PP#123), CDAO (PP#125), and geiger (PP#27). (The identifiers that are referenced here are primary keys in the data tables that we provide at https://nescent.github.io/community-and-code/data/ - specifically, the items listed here are from the *project_products* table, hence their primary keys have the "PP" prefix.)

Of the repositories developed for the OpenTree hackathon, only two remained active after the hackathon, both from the team that developed library wrappers for OpenTree services. One of the active repositories (PP#140) has an innovative test system that uses the same interface to test Python, Ruby and R libraries. The other is for the R library rotl, which we discuss further in the next section.

In the case of incremental additions to existing codebases, the impact of any new code is difficult to judge, unless the code adds distinctly new features whose use can be tracked. For instance, the "Integrating Ontologies" team mentioned above added useful features to a previous hackathon product, an XSLT stylesheet for translating between NeXML and CDAO. This enhanced version is still in use in the production version of TreeBASE to provide trees in an RDF-XML format.

Hackathon teams sometimes produce documentation, though much less commonly than code. Sometimes this took the form of screencasts illustrating prototypes. Of two Phylotastic screencasts, one of them (PP#115) has received 420 views, and another (PP#130) has received 200 views (at time of writing). Perhaps a more useful documentation product is the "phylogenetics" task view (event product EP#15) on the Comprehensive R Archive Network (CRAN), which provides a concise synopsis of available R packages for phylogenetics. This documentation continues to be updated (most recently 2017-04-09), but we have no way of knowing how frequently it is used.

Sometimes, the main product of a team is a design or schema. Team #52 ("skelesim") from the R popgen hackathon aimed to integrate several different simulation packages in a common framework: this goal proved far too ambitious, and the group had only a design when

Table 3. Hackathon outcomes and possible impact metrics.

Outcome	Definition	Example	Output metrics	Impact metrics
Code	Source code to be compiled or executed	OpenTree API library code	Number of lines	Number of reverse depends, uses, users, mentions
Documentation	Instructions and examples	R vignettes	Number of lines	Number of users, mentions
Designs, Schemas	Partially or fully encoded products that are not executed	Phylo-querying design	Number of lines	Number of users, mentions
Installations	Installing and running software	Live Phylotastic pruner demo	Number of functions	Number of users, mentions
Data Products	Aggregation, QA, etc., resulting in value-added data product	Formal annotation of high-value trees	File size, number of data	Number of downloads, mentions
Training	New or increased proficiency in a skill	Learning GitHub	Proficiency	(output metric for using proficiency)
Experience	New or increased exposure to circumstances	Team programming	NA	NA
Networking	New or enhanced professional connections	NA	Number of new or strengthened connections	Number of collaborations
Awareness	Recognition of resources, challenges, practices	Name-resolution is a major issue	Number of newly recognized technologies, challenges, best practices	Uses of newly recognized technology
Posters	Graphic presentations for meetings	R popgen poster	Number of posters, number of times presented	Number of mentions
Blogs	Non-reviewed web publications	TreeForAll blogs	Number of blogs	Number of mentions
Presentations	Conference presentations	DateLife iEvoBio presentation	Number of presentations, number of times presented	Number of mentions
Collaborations	Commitment to pursue shared goal	rotl collaboration	Number of collaborations	(output metric for collaboration)
Manuscripts	Manuscripts for publication	Phylotastic	Number of manuscripts, number accepted	Number of citations, mentions
Proposals	Proposals for funding further work	Phylotastic proposal	Number of proposals, number accepted	(output metric for funding)

the hackathon ended (PP#112). Another example of a team tackling a difficult challenge was the work of Team 20 on "phyloreferencing", essentially a topological query language for trees. This work was important in subsequently securing major funding (PP#139).

Some additional kinds of products are rather infrequent. A unique tangible outcome of the first R hackathon was the development of an email list that is still in use today, the r-sig-phylo list (EP#4), which we mention further below. Though it sometimes happens (e.g., EP#9), hackathon teams rarely provide a public-facing demonstration, because those require a high level of completion and extra effort. One hackathon team worked on a data product, consisting of an annotated set of high-value phylogenetic trees (PP#6). The challenge was to develop a completely machine-readable scheme of annotation based on available ontologies. The trees were not used subsequently.

Tangible follow-up outcomes and their impacts

Beyond the direct code produced at the hackathon, tangible outcomes may continue to develop after the event is over. In fact, most of the time that a hackathon product has a major impact, this is due to follow-up work by participants. The most common of these are (1) demonstrations and production code, (2) communications such as blogs or meeting presentations, (3) manuscripts for publication, and (4) proposals for funding to support further work.

Sometimes an individual participant continues working after an event, either to finish up a specific product, or simply acting on a burst of enthusiasm. An example of the latter would be an enormous spike of 238 commits to DendroPy by a single individual in a month-long period beginning a week before event 6, and continuing two weeks after (PP#134). We also identified some cases in which individuals developed a formal communication, such as

a blog series (e.g., EP#11) or a meeting presentation (e.g., EP#16, EP#17).

More commonly, follow-up activities emerge in the context of a group commitment to continue working together. The "PhyloGeoTastic" team did not finish their implementation at the hackathon, but this was completed afterwards so that a live demonstration (PP#4) would be available (although this demo has subsequently gone offline). Two of the events produced published reports that included all of the hackathon participants as authors (EP#1, EP#10;[8,9]). In both cases, the process of writing and submitting the articles took many months, and was driven and managed by the organizers, reflecting an uneven commitment, with some individuals contributing much more than others. The three participants at the OpenTree hackathon who worked on the R library, now called "rotl", developed this product mostly after the hackathon (with hundreds of commits), leading to a publication[38], and a package that has already been used in a subsequent scientific study[39].

Occasionally, participants pursue funding for more extensive follow-ups. Several projects led to Google Summer-of-Code proposals (EP#12, EP#13, PP#137), two of which were funded. The "skelesim" group at the R popgen hackathon later wrote a proposal (PP#129) that won funding for a four-day meeting to continue their work several months later. Two participants in the "phyloreferencing" group at the VoCamp eventually wrote a grant proposal (PP#139) and secured three years of funding for that project. The two Phylotastic hackathons also led to a successful proposal for major funding (EP#8). So, if we look at hackathons as proposal germinators, this is rather high impact. The total award amount for the two National Science Foundation grants is approximately $7.5 M. By comparison, the total amount spent on the nine NESCent hackathons was roughly $250,000. Of course, one cannot calculate a return-on-investment from these two numbers alone, because it does not take into account the significant amount of post-hackathon work required to write a proposal. However, if a grant proposal typically results from three modestly paid academics working quarter-time for three months, and there are not a large number of failed proposals that we are not counting (and we know of no failed major proposals based on NESCent hackathon products), this does not change the overall impression that hackathons are a good investment.

Intangible outcomes, immediate or follow-up

Other authors have pointed out that hackathons are highly social events that provide opportunities to build relationships[19,28] and experience excitement around shared motivations[16,40]. However, such intangible outcomes are difficult to document. In some cases, an intangible outcome is apparent because it has a tangible impact. For example, while the r-sig-phylo mailing list (EP#4) was a direct outcome of the Comparative Methods in R hackathon, the mere existence of a list did not guarantee that this list would be used subsequently, nor that it would garner any new subscribers beyond the initial set of 28 participants. However, eight years later, the mailing list now has 1155 subscribers (as of December 5th, 2016) and generates approximately thirty to sixty messages per month. From this, we would argue that the hackathon helped to nurture a community of practice as an intangible outcome.

Another example of a case where a hackathon helped to foster a new community as an intangible outcome is the follow-up on team #52, whose six members, most of whom had not worked together before, had become sufficiently motivated to obtain funding for a second face-to-face meeting of four days (mentioned above), and then to meet virtually on a biweekly basis for eight months, in order to finish a project and submit a manuscript on it. Likewise, two of the authors of the present manuscript (AS, EP) are leaders of an ongoing Phylotastic project, and several hackathon participants are consultants. Yet, of the many code repositories developed by teams participating in the two Phylotastic hackathons, only one code repository remains in active development (PP#2 for DateLife, part of the funded project), while a second repository is maintained for providing web content. The continuity between the hackathon and the current project is primarily a continuity of people, plans, excitement, and working relationships, not a continuity of code. In a somewhat similar way, the web-services interface to TreeBASE, written at the third hackathon (see PP#37) was not used in TreeBASE, but the main author later wrote a production version based on the initial implementation. The intangible outcome in this case was the knowledge and the confidence that a particular problem could definitely be solved.

We would argue from such examples that the experience of a hackathon results in intangible outcomes that sometimes yield tangible benefits. The intangible outcomes are various forms of technical learning, developing a shared awareness (e.g. of what is technically possible), and involve building new relationships. Participants seem to understand this: Briscoe *et al.*[28] report survey results indicating that the top two reasons for participation in hackathons are "learning" and "networking." Again, we cannot document these intangible outcomes in a direct and rigorous way, but we suggest that some of the following are worth considering:

Technology learning: NESCent hackathons often relied on a recommended set of assistive technologies. Many participants learned these technologies for the first time; e.g. obtaining GitHub accounts and learning how to use GitHub. They also learned about specific resources (e.g. code libraries) while participating.

Exposure to best practices: In many cases, hackathons provided scientific programmers with critical exposure to best practices widely accepted among professional programmers, such as using collaborative versioning systems, writing documentation, and running automated unit tests.

Opportunity to learn: In some cases, the goal of a group was simply to learn by doing. For example, the "integrating ontologies" group at the VoCamp did not have a functional goal in mind, but aimed to do hands-on work in order to learn how to build bridges between ontologies.

Team programming experience: Obviously a hackathon provides the actual experience of coding, but the team-based aspect of this experience is often novel: scientific programmers frequently work alone. For many, the chance to discuss designs and develop code with a team or as a pair ("pair programming") is a rare opportunity.

Awareness of technical challenges and opportunities: Discussion and information-sharing often had the effect of promoting a shared understanding of technical challenges and opportunities. This is vital in a technology landscape that is constantly changing, especially in the evolutionary informatics community, which is as a small and dispersed community.

Lessons learned

In this paper, we describe nine hackathons that we co-organized and participated in, in varying teams, over the course of roughly a decade. After the last of these hackathons, we re-convened at NESCent in a separate meeting to discuss and summarize our experiences. Over the course of this meeting, during our email discussions afterwards, and during the writing of this paper, we developed a synthesis of "lessons learned" that we all agree with as being key in organizing a NESCent hackathon. In this section we discuss these lessons.

Lesson 1: Choose a clear yet flexible theme. In our experience, a well chosen theme: (1) leverages the skills and interests of likely participants in such a way that the projects that emerge will serve the goals of the hackathon as identified during the initial scoping (and will align with the interests of sponsors); and (2) allows abundant flexibility for participants to exercise creativity and maximize the value of their participation, including their desire for learning and networking[28]. We are inspired by the OpenSpace philosophy[41] that a theme "must have the capacity to inspire participation by being specific enough to indicate the direction, while possessing sufficient openness to allow for the imagination of the group to take over". The importance of having a well-defined problem or theme that is communicated effectively to participants has also been stressed by Mohajer Soltani *et al.*[18]. Others have suggested that the hackathon should balance a sponsor's desire for tangible outcomes with the participants desire to learn[42].

In our experience with organizing hackathons, the scope of the hackathon typically emerges after organizers have reflected on community needs. This sometimes involves pre-event discussions with participants, like the use-case driven approach in 8,33. The scope typically has both a technological and a thematic aspect. For instance, in the case of the two R hackathons, the technological limitation was the use of R, and the domain of application was either population genetics or comparative phylogenetic methods.

Our choices of scope were not always ideal. In the case of the VoCamp, the scope was loosely focused on the intersection of evolutionary biology with "ontologies and controlled vocabularies". With a theme that is too broad, the pre-pitching discussion is diffuse and there is little reason to value one idea over another. This makes it less likely that strong teams will emerge. For an entirely different reason, the second Phylotastic was also not as successful. The first Phylotastic hackathon took a good idea (to create an ecosystem of web services that deliver time-calibrated subsets of the Tree of Life) and turned it into a prototype, which created a large amount of excitement in the community. The second Phylotastic hackathon had essentially the same theme, which meant that for projects to be successful, they had to go beyond prototyping, without relying on a drawn-out process of analysis or design that had not yet happened.

Lesson 2: Build the right leadership team. Leadership team members are often selected among more established and senior researchers. The benefit is that they have a greater level of awareness of the community and may provide better guidance in the scoping of the problem and in the identification of effective participants. On the other hand, more senior researchers tend to have an extensive agenda of commitments, which detracts from the dedication required by the organization of a hackathon event that is highly focused, intense over an extended period of time, and guided by rigid deadlines. LT members need to be available for regular meetings (e.g. weekly or bi-weekly teleconferences), and participate in the preparatory activities (e.g. attending pre-hackathon working group meetings; preparing, disseminating and evaluating applications).

Commitments of LT members tend to change rapidly, leading to shifts in focus and in the level of engagement. There have been instances where LT members have had to abandon the team due to a sudden lack of time and availability. Major shifts in the LT composition may endanger the success of the event. Just as the initial success of the hackathon event is dependent on the time and effort dedicated by members of the LT, they are also often the individuals that organize the post-hackathon activities to summarize results, guide follow-up efforts (e.g. development of manuscripts that present the achievements of the hackathon projects), and ensure that the hackathon outcomes are made fully available to the broader community.

Lesson 3: Pre-select assistive technologies. Many online platforms are available to assist in communication and collaboration. These include text chat, teleconferencing and videoconferencing; collaborative document editing platforms; issue trackers and project management tools; and source code revision control systems. It is a good idea to pre-select certain preferred technologies from amongst these, and commit to them. Allowing multiple technologies reduces the chances of effective coordination among participants during the event, and also impedes any post-event attempts to create a cohesive record or to track outcomes.

In our experience, ideal assistive technologies are ones that allow you to track activity and outcomes so that they feed into a system of record-keeping and results-tracking. As has been noted in more formal systematic reviews and meta-analyses[43], it is not obvious how much data is missing from the literature until one attempts to collect the data. For many hackathons, we used wikis for open document planning and note taking. This resulted in a rich historical record; yet we still find that basic data about the hackathons can be hard to compile because there was not a clear plan for gathering, organizing and preserving information.

Encouraging the common use of source code revision control systems such as GitHub offers many opportunities to access information about (1) contributions made by participants, such as the extent

of their usage of such platforms before the event, during, and after (2) the development of the source code repositories worked on, and (3) the dynamics of collaborations, for example looking at the degree to which hackathon participants worked on the same repositories before, during, and after the event.

Lesson 4: Diversify and grow the community. Our experience with hackathons is in academic settings, and so our participants have been a mix of faculty, postdocs, students, and research staff. Research staff are less likely to be able to participate in post-hackathon commitments after the hackathon is over because of their busier schedules; they are also less likely to generate a career benefit from a product. Conversely, postdocs and students may be able to engage more fully, including in the preparatory and follow-up stages, but will benefit from the presence of more senior faculty that might provide informal mentoring opportunities[2,42,44]. Hence, like others (e.g. 18,40), we recognize the importance of diversity in participant competences and career stages, and made efforts to balance diversity in this respect.

In addition to assembling participants with diverse levels of expertise, we also made a conscious effort to bring together and benefit from international participants, as well as participants from traditionally underrepresented groups. One strategy to increase diversity is to pay attention to the language used in recruitment materials, bearing in mind that women often undervalue their skills relative to men[45]. Thus, we avoided announcements that seemed to set a highly restrictive standard of technical skill or domain knowledge; i.e. appealing to "power users" would be a mistake (and appealing to "gurus" or "wizards" would be worse). However, our main strategy for increasing diversity was targeted invitation: we identified qualified participants who could increase the diversity of the event, and personally invited them to apply. Women and scientists from minority groups in senior positions are often good sources for providing names of other women and scientists from minority groups in junior positions. It also helps to have a diverse organizing team. In practice, we assembled a list of candidates, and split the task of writing personal invitations among the leadership team. Whereas our open call (distributed electronically) reached thousands of people and generated only a few dozen applications at most, we estimated that applications were received from about 1 out of 2 people personally invited to apply by a hackathon organizer. In our experience, applicants recruited in this manner have similar qualifications to other applicants, and have roughly the same chance of being accepted.

Direct invitation to a hackathon serves not only to increase diversity, but also to target expertise. However, we found it to be too limiting as a general strategy: choosing participants from an open pool is important if one of the goals of the hackathon is to grow the community, whereas invitation-only hackathons (e.g. the BioHackathons[23–26] organized by DBCLS, Japan) have a danger of ossifying patterns of inclusion and exclusion.

Lesson 5: Engage participants early on. A well-organized hackathon includes sufficient pre-event engagement with participants so that they can hit the ground running on day one. A number of topics need to be addressed; a well-defined theme needs to be effectively communicated to the participants[18]; there needs to be group consensus on objectives and their domain-specific context[10]; and any assistive technologies need to be chosen and their requirements assessed. Pre-event engagement intends to ensure that participants are well-prepared in practical terms. For example, this includes them having practiced with new technologies[19] - and having signed up for such technologies ahead of time, if need be. However, beyond such practical preparation, they should also have mentally worked themselves up for "invested participation"[28] in the event.

The need for sufficient time prior to the event is emphasized by Christopherson *et al.*[10], who describes two successive hackathons with vastly different amounts of time to engage participants and develop ideas prior to the respective events: "This time crunch added undue pressure on the team, and some participants reported that this made it more difficult to achieve synergy as quickly as expected. It ultimately resulted in less working code . . . and contributed to lower reported satisfaction."

At the NESCent hackathons, most of the pre-event engagement was on a mailing list that participants were subscribed to as soon as possible. To foster community engagement, we used the same mailing list and simply added new members for each hackathon. We also experimented with real-time communication prior to the events, using videoconferencing (Google Hangouts, prior to Phylotastic 1). Providing opportunities for engagement can be effective even if only a minority of participants are involved: the ones who feel the greatest need to prepare and to learn more about what will happen at the hackathon are the ones most likely to participate.

Lesson 6: Be welcoming and encouraging. As many have pointed out,[16,19,40,46], hackathons are highly social events where success depends on what Briscoe and Mulligan call "invested participation"[28]. That is, participants must feel invested personally in the event. Yet, hackathons have earned a reputation as unwelcoming events catering to insiders and to men. To remedy this, we have made conscious efforts to communicate in ways that are welcoming and encouraging, and to manage the hackathon event in ways that are welcoming and encouraging.

First, we designed recruitment materials to appeal to a wide audience, avoiding highly technical language except where absolutely necessary. We explicitly specified non-programmer roles (e.g. "domain expert", "use-case consultant"), and avoided implicitly equating participants with programmers (e.g. we did not refer to them as "programmers" or "coders").

Second, we made it a practice, prior to the event, to reach out personally to individuals who were not already part of our professional network (typically one to two-thirds of the participants were new to us). In most cases this was as simple as an organizer writing a brief email thanking the individual for applying and offering a statement of encouragement about participating in the upcoming event. During the event, there were many opportunities to improve participation by making people feel welcome, e.g. expressing appreciation for opinions and suggestions that were brought forward by newcomers. To ensure that everyone can participate fully,

novice participants were given permission to join a team simply to learn and assist, even if they did not have sufficient technical or domain knowledge to be a key contributor. One of the ways we encouraged this practice, was to tell participants that, after teams formed and work started, we wanted everyone to be "either learning or doing".

During team formation, facilitators may intervene to discourage teams from unintentionally closing ranks around a pitch (some participants will commit early to a pitch and begin deep technical discussions, sometimes with their backs to everyone else, which discourages others from approaching or getting involved). When organizers acted as discussion facilitators, they would model the process of asking non-negative, open-ended questions; rather than "Isn't that out of scope?", they would ask "What are some ways that this idea aligns with our goals of...?".

Lesson 7: Minimize remote participation. The technological possibilities for remote and asynchronous collaboration make it seem superficially attractive for hackathon organizers to expand the scope of the event by supporting remote participants. However, remote participation is not without cost, and is typically considerably less effective than direct participation. Most potential team-mates at a hackathon have not collaborated before. The face-to-face and real-time nature of hackathons allows for considerable transfer of information that turns out to be quite frustrating to achieve through remote communication, costing extra time to deal with lossy communication and lacking in-person dynamics like whiteboarding or looking over a shoulder[10].

Organizers should therefore consider in advance whether they will support remote participation (strategies for doing so are described at https://nescent.github.io/community-and-code/doc/concise_guide/remotes/). Allowing remote participation has the advantage of reaching more people in the community, and expanding the productive capacity of the hackathon, but it carries a risk of frustration, and comes at a predictable cost as there is a burden to supporting remote participation, e.g. an increased demand for on-site participants to adhere closely to a fixed schedule.

We explored various means of including remote participants. In one case, we made an arrangement for a satellite hackathon to be held in parallel by a small group on the west coast (NESCent is on the east coast). The westcoast participants were all from the same research group: they formed a single team that contributed importantly to the hackathon. This allowed a significant expansion of the scope of the hackathon, at no monetary cost, and with little trouble.

Single individuals also participated remotely in NESCent hackathons on numerous occasions, with uneven success. The factors that seemed to contribute to success included their level of previous experience with hackathons and remote collaboration, a commitment to avoid local distractions, and a clear sense of where to fit into a team project. Perhaps most importantly, in all successful cases, the remote person was already part of the community and had collaborators on site. By contrast, remote participation is not an advisable way to include new people.

We recommend a buddy system where each remote participant is paired with an on-site participant who maintains a video connection throughout the meeting, and serves as a conduit for communication at team work sessions and plenary sessions. Sticking to specific communication technologies is also critical; if the in-person team changes technology halfway through, the remote participants may quickly become lost and forgotten.

Lesson 8: Manage the team formation process. The coalescence of participants into teams is a critical step. At some hackathons discussed in the literature, the teams were fully pre-specified, with no obvious team formation process during the event[16], while other hackathons were organized around, for example, student projects[32] or the desire to learn a new technology[44]. At the NESCent hackathons, we emphasized self-organization. At the first hackathon in 2006, this self-organization was guided firstly by use cases that had been decided upon by the participants prior to the event, and secondly by participants' existing connections to open source software projects, such as the Bio* toolkits. However, this may have impeded the building of new connections. At later hackathons, the group formation process was deferred to the first day of the event itself.

Several authors have discussed the social nature of the team-formation process (e.g. Jones *et al.*[46]), and we have also sought to promote this. We did so by arranging for final team formation to occur by a facilitated self-organizing process, described in more detail at https://nescent.github.io/community-and-code/doc/concise_guide/managing/. During the final stage of this process, participants pitch hackathon activities to each other using whiteboards or flipboards, and teams coalesce around pitches in a manner akin to OpenSpace team coalescence (e.g. as in Mulholland and Meredig[12]). It is important to precede this stage with an opportunity for the group to discuss the relevance, importance, and chances of success of pitches, to ensure that weak points are addressed.

Lesson 9. Manage expectations for follow up. When hackathon teams are working energetically, organizers and team members may have enthusiastic discussions of follow-ups, yet when the hackathon ends, team dynamics and energy often dissipate rapidly as team members return to other responsibilities, resulting in little follow-up (e.g. 21). After all, the nature of hackathons is that we steal talented people from their day jobs for a limited time, and so a team dynamic during the event is unlikely to persist beyond the face-to-face conditions that fostered the team (though this does occasionally happen, e.g. 2).

We therefore adopted two strategies to manage expectations for follow up. The first strategy involved accepting that, because of the low prospects for follow-up, organizers should instruct participants to focus on producing tangible products within the space of the hackathon, with the expectation that tasks unfinished on the last day would never be finished. The second strategy involved encouraging commitment to a follow-up program to build on successful projects (an example of this approach is in 32). In several cases, NESCent hackathon projects have provided proofs-of-concept and specifications that were important for obtaining funding for further development. In two cases, our hackathons resulted directly in a scholarly publication[8,9], with additional examples found elsewhere (e.g. 30).

Because the cases of successful follow-ups are small in number, it is hard to make generalizations. However, it seems obvious that the potential for follow-up increases when a hackathon project aligns with the interests of participants, and with a leader who has the time to manage a follow-up effort. The more junior participants may be more driven to pursue a project after an event because the outcome may have a larger career impact. Thus, to achieve tangible working products and manage followup, one should consider two things. First, whether deliverables can be achieved in the time allotted for the hackathon, and second, whether choosing participants who may lack certain skills or experience would serve the overall goal better by their likelihood of being able to dedicate extra time after the event to follow-up and complete the project. The success of this latter strategy is further influenced by the ability to coalesce a wider community around activities performed at a hackathon. The development of open data repositories and the contribution to widely accessed code repositories are aspects that facilitate the community "buy-in" and enable long-term sustainability of hackathon products. An example of this is represented by the contributions to the NeXML code base achieved during the Database Interoperability hackathon.

Conclusions

We have provided systematic information on events, participants, teams, projects, and outcomes pertaining to the nine NESCent hackathons that took place from 2006 to 2015. NESCent hackathons represent a unique form of participant-driven software development meeting. The NESCent model was designed, not only to stimulate software development and to provide training and experience to participants, but to nurture a larger community of practice so that members develop a shared awareness of best practices, available resources, and strategic challenges. To allow others to use this model, we have developed detailed guidance and sample materials for planning, advertising, recruiting, and facilitation.

The impacts of hackathons depend on tangible and intangible outcomes. The most obvious tangible outcome of a hackathon is computer code. Some hackathon teams made incremental additions of code or documentation to pre-existing (production) codebases, but most produced standalone products such as prototypes, draft standards, or designs. Standalone products are rarely used or maintained after the hackathon ends, but may have downstream impacts as inspiration or proof-of-concept. To date, two NESCent hackathon projects have led to major NSF funding for the development of production systems. In addition, NESCent hackathons have led rather directly to 4 publications, along with various posters, blogs, websites, and presentations. The intangible impacts of NESCent hackathons, which are perhaps more important, are much more difficult to track. We have described some indications of positive impacts of hackathon-associated training, networking, and community-building, but we can draw no firm conclusions in these areas.

Without systematic information on other types of hackathons, one cannot draw conclusions on the effectiveness of the NESCent model relative to other types of hackathon. Indeed, a direct comparison with other hackathon types that are designed with different aims may not be appropriate, given the specialized aims of the NESCent model to serve a geographically dispersed academic community. Nevertheless, we hope that the systematic information provided here will lay the foundation for future research on the effectiveness of participant-driven meetings.

Data availability

To accompany this paper, we have developed a website that contains a *Concise Guide* for organizing NESCent hackathons here: https://nescent.github.io/community-and-code/doc/

In addition, we collated a dataset itemizing all events; their participants; the projects worked on at each event; and the outcomes, both at the level of individual projects and at the level of events. We make this dataset available for download as machine readable, tab-separated tables, here: https://nescent.github.io/community-and-code/data/

These tables summarize, in structured form, data that were previously spread out over the different websites and wikis that were used for each hackathon, and that since may have gone offline (or moved) subsequent to NESCent's closing. As such, all data were, and are, in the public domain.

Author contributions

AB: data entry, data management; ms writing. AS: data entry, data management; literature analysis; ms writing; guidelines writing. EP: data entry, data management; ms writing. HL: data entry, data management; ms writing. KC: data entry, data management; ms writing. MSR: ms and supplement writing. RV: data entry, data management; literature analysis; ms writing. SO: ms writing.

Competing interests

No competing interests were disclosed.

Grant information

The National Evolutionary Synthesis Center (NSF #EF-0905606) extended its support for the HIP (Hackathons, Interoperability, Phylogenies) working group by funding the authors to gather at NESCent in March of 2015 to lay the groundwork for this project.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

We are grateful to the organizations and projects that supported the hackathons described here, including NESCent, the iPlant Collaborative, the Biodiversity Synthesis Center of the Encyclopedia of Life (BioSync), Biodiversity Information Standards (TDWG) and the OpenTree project. The identification of any specific commercial products is for the purpose of specifying a protocol, and does not imply a recommendation or endorsement by the National Institute of Standards and Technology.

References

1. Caron B: **Getting a handle on community.** *Figshare.* Blog Post; 2015.
Publisher Full Text
2. Busby B, Dillman A, Simpson CL, *et al.*: **Building Genomic Analysis Pipelines in a Hackathon Setting with Bioinformatician Teams: DNA-seq, Epigenomics, Metagenomics and RNA-seq.** *bioRxiv.* 2015.
Publisher Full Text
3. Almirall E, Lee M, Majchrzak A: **Open innovation requires integrated competition-community ecosystems: Lessons learned from civic open innovation.** *Bus Horiz.* 2014; **57**(3): 391–400.
Publisher Full Text
4. Aungst TD: **Using a hackathon for interprofessional health education opportunities.** *J Med Syst.* 2015; **39**(5): 60.
PubMed Abstract | Publisher Full Text
5. DePasse JW, Carroll R, Ippolito A, *et al.*: **Less noise, more hacking: how to deploy principles from MIT's hacking medicine to accelerate health care.** *Int J Technol Assess Health Care.* 2014; **30**(3): 260–4.
PubMed Abstract | Publisher Full Text
6. McArthu K, Lainchbury H, Horn D: **Open Data Hackathon: How to Guide.** Web Page, 2012.
Reference Source
7. Murby R: **How to throw the perfect hackathon**. Blog Post, 2014.
Reference Source
8. Lapp H, Bala S, Balhoff JP, *et al.*: **The 2006 NESCent Phyloinformatics Hackathon: A field report.** *Evol Bioinform.* 2007; **3**: 287–296.
Free Full Text
9. Stoltzfus A, Lapp H, Matasci N, *et al.*: **Phylotastic! Making tree-of-life knowledge accessible, reusable and convenient.** *BMC Bioinformatics.* 2013; **14**: 158.
PubMed Abstract | Publisher Full Text | Free Full Text
10. Christopherson L, Idaszak R, Ahalt S: **Developing Scientific Software through the Open Community Engagement Process.** *Figshare.* 2015.
Publisher Full Text
11. Tanenabaum K, Tanenbaum JG, Williams AM, *et al.*: **Critical making hackathon: situated hacking, surveillance and big data proposal.** In *CHI '14 Extended Abstracts on Human Factors in Computing Systems.* 2014.
Publisher Full Text
12. Mulholland G, Meredig B: **Hackathon aims to solve materials problems.** *MRS Bull.* 2015; **40**(4): 366–370.
Publisher Full Text
13. Kuchinskas S: **How to run a winning hackathon.** *Forbes.* 2014.
Reference Source
14. Johnson P, Robinson P: **Civic hackathons: Innovation, procurement, or civic engagement?** *Rev Policy Res.* 2014; **31**(4): 349–357.
Publisher Full Text
15. Burnham K: **Inside facebook's hackathons: 5 tips for hosting your own.** *CIO.* 2012.
Reference Source
16. Raatikainen M, Komssi M, Dal Bianco V, *et al.*: **Industrial experiences of organizing a hackathon to assess a device-centric cloud ecosystem.** In *Computer Software and Applications Conference (COMPSAC), 2013 IEEE 37th Annual.* 2013; 790–799.
Publisher Full Text
17. Bond RR, Mulvenna MD, Finlay DD, *et al.*: **Multi-faceted informatics system for digitising and streamlining the reablement care model.** *J Biomed Inform.* 2015; **56**: 30–41.
PubMed Abstract | Publisher Full Text
18. Soltani PM, Pessi K, Ahlin K, *et al.*: **Hackathon: A method for digital innovative success: A comparative descriptive study.** In *Proceedings of the 8th European Conference on IS Management and Evaluation.* Academic Conference and Publishing International, Ltd. 2014; 367–373.
Reference Source
19. Trainer EH, Chaihirunkarn C, Kalyanasundaram A, *et al.*: **Community code engagements: Summer of code and hackathons for community building in scientific software.** In *Proceedings of the 18th International Conference on Supporting Group Work.* ACM Press. 2014; 111–121.
Publisher Full Text
20. Calco M, Veeck A: **The markathon: Adapting the hackathon model for an introductory marketing class project.** *Mark Educ Rev.* 2015; **25**(1): 33–38.
Publisher Full Text
21. Knight Foundation: **Four ideas for the future of hackathons.** Blog Post. 2012.
Reference Source
22. Komssi M, Pichlis D, Raatikainen M, *et al.*: **What are hackathons for?** *Software, IEEE.* 2015; **32**(5): 60–67.
Publisher Full Text
23. Katayama T, Arakawa K, Nakao M, *et al.*: **The DBCLS BioHackathon: standardization and interoperability for bioinformatics web services and workflows. The DBCLS BioHackathon Consortium.** *J Biomed Semantics.* 2010; **1**(1): 8.
PubMed Abstract | Publisher Full Text | Free Full Text
24. Katayama T, Wilkinson MD, Vos R, *et al.*: **The 2nd DBCLS BioHackathon: interoperable bioinformatics Web services for integrated applications.** *J Biomed Semantics.* 2011; **2**: 4.
PubMed Abstract | Publisher Full Text | Free Full Text
25. Katayama T, Wilkinson MD, Micklem G, *et al.*: **The 3rd DBCLS BioHackathon: improving life science data integration with Semantic Web technologies.** *J Biomed Semantics.* 2013; **4**(1): 6.
PubMed Abstract | Publisher Full Text | Free Full Text
26. Katayama T, Wilkinson MD, Aoki-Kinoshita KF, *et al.*: **BioHackathon series in 2011 and 2012: penetration of ontology and linked data in life science domains.** *J Biomed Semantics.* 2014; **5**(1): 5.
PubMed Abstract | Publisher Full Text | Free Full Text
27. Möller S, Afgan E, Banck M, *et al.*: **Community-driven development for computational biology at sprints, hackathons and codefests.** *BMC Bioinformatics.* 2014; **15**(Suppl 14): S7.
PubMed Abstract | Publisher Full Text | Free Full Text
28. Briscoe G, Mulligan C: **Digital innovation: The hackathon phenomenon.** Report, CreativeWorks London, 2014.
Reference Source
29. Correia NN, Tanaka A: **Prototyping audiovisual performance tools: A hackathon approach.** In *Proceedings of the international conference on New Interfaces for Musical Expression.* 2015.
Reference Source
30. Fafalios P, Papadakos P: **Theophrastus: On Demand and Real-Time Automatic Annotation and Exploration of (Web) Documents using Open Linked Data.** *Web Semant.* Elsevier, 2014; **29**: 31–38.
Publisher Full Text
31. Groen D, Calderhead B: **Science hackathons for developing interdisciplinary research and collaborations.** *eLife.* 2015; **4**: e09944.
PubMed Abstract | Publisher Full Text | Free Full Text
32. Hecht BA, Werner J, Raskar R, *et al.*: **The kumbhthon technical hackathon for nashik: A model for stem education and social entrepreneurship.** In *Integrated STEM Education Conference (ISEC), 2014 IEEE.* 2014; 1–5.
Publisher Full Text
33. Vos RA, Biserkov JV, Balech B, *et al.*: **Enriched biodiversity data as a resource and service.** *Biodivers Data J.* 2014; (2): e1125.
PubMed Abstract | Publisher Full Text | Free Full Text
34. Souza S: **Lessons learnt from a public-private big data hackathon.** Report, Big Innovation Centre, 2013.
Reference Source
35. Lave J, Wenger E: **Situated Learning. Legitimate peripheral participation.** University of Cambridge Press, Cambridge, 1991.
Reference Source
36. Wenger-Trayner E, Wenger-Trayner B: **Communities of Practice: a Brief Introduction.** 2015.
Reference Source
37. Vos RA, Caravas J, Hartmann K, *et al.*: **BIO::Phylo-phyloinformatic analysis using perl.** *BMC Bioinformatics.* 2011; **12**: 63.
PubMed Abstract | Publisher Full Text | Free Full Text
38. Michonneau F, Brown J, Winter D: **rotl, an R package to interact with the Open Tree of Life data.** *Methods Ecol Evol.* 2016; **7**(12): 1476–1481.
Publisher Full Text
39. Brandvain Y, Coop G: **Sperm should evolve to make female meiosis fair.** *Evolution.* 2015; **69**(4): 1004–14.
PubMed Abstract | Publisher Full Text
40. Zapico JL, Pargman D, Ebner H, *et al.*: **Hacking sustainability: Broadening participation through green hackathons.** In *Fourth International Symposium on End-User Development.* 2013.
Reference Source
41. Owen H: **A brief user's guide to open space technology.** *Nd,* 2008.
Reference Source
42. Linnell N, Figueira S, Chintala N, *et al.*: **Hack for the homeless: A humanitarian technology hackathon.** In *Global Humanitarian Technology Conference (GHTC), 2014 IEEE.* 2014; 577–584.
Publisher Full Text
43. Koricheva J, Gurevitch J, Mengersen K: **Handbook of Meta-analysis in Ecology and Evolution.** Princeton University Press, Princeton, 2013.
Publisher Full Text
44. Mtsweni J, Abdullah H: **Stimulating and maintaining students' interest in computer science using the hackathon model.** *The Independent Journal of Teaching and Learning.* 2015; **10**: 85–97.
Reference Source
45. Nafus D, Leach J, Krieger B: **Gender: Integrated report of findings.** Report, University of Cambridge, 2006.
Reference Source
46. Jones GM, Semel B, Le A: **There's no rules. it's a hackathon: Negotiating commitment in a context of volatile sociality.** *J Linguist Anthropol.* 2015; **25**(3): 322–345.
Publisher Full Text

Open Peer Review

Current Peer Review Status: ✔ ?

Version 1

Reviewer Report 21 June 2017

https://doi.org/10.5256/f1000research.12340.r23275

© **2017 Neylon C.** This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

?

Cameron Neylon
Centre for Culture and Technology, Curtin University, Perth, WA, Australia

This article provides a valuable introduction and "lessons" perspective based on a decade of Hackathons based at NESCent. This provides a valuable perspective on the conduct and impact of this approach in an academic setting and provides a good bases for developing assessment approaches for the value of Hackathons in academic settings.

I concur with the first referee that more background would enhance the value of the paper as it stands.

My primary concern is with the way "intangible" outcomes are presented as "naturally hard to assess". I'm not convinced firstly with the use of the word "intangible" in this context. In many ways the primary outputs of the hacks are "intangible" code, and things described as intangible (community building, enhanced training/skills) are in some ways quite tangible. I can't readily offer a replacement term, which suggests to me that actually framing this distinction in a different way might be valuable.

I would suggest considering using the language of research impact assessment of "outputs", "outcomes", and "impacts". The "tangible outcomes" described here are mostly "outputs", specific objects that are the result of the work. The "intangible outcomes" are largely "outcomes", that is effects and consequences that result from the outputs. Finally the goal of the paper seems to be to work towards ways of describing "impacts", measurable changes in the world that result from the project.

I feel this matters because there is a tendency in the article to discount the possibility of evaluating these "intangible" outcomes, while at the same time there is a recognition that they may be the most important outcomes and impacts of the hack. In turn there *are* ways to measure these: user surveys, network analysis (pre and post), case studies (as is explored a little here), but these approaches require a degree of preparation and analysis beyond that reported here.

I think the article would therefore benefit from a close look at what approaches might be deployed

to measure these outcomes and impacts, and what the challenges of doing that might be. In my view this would enhance the paper significantly, particularly in its goal of starting a conversation on how best to evaluate and communicate the success of these events.

More specific comments:

In the Abstract: “Intangible outcomes could not be assessed objectively…” Suggest rewording this in any case. Intangible outcomes can of course be assessed objectively, just perhaps not directly. Regardless of whether you choose to make the other changes I'd reword this to something like "The less tangible [direct?] outcomes of the events are harder to track and measure, but may be amongst the most important".

Introduction: There is an online literature critical of (particularly commercial) hackathons which might be worth touching on as a negative side, even if it is not peer reviewed. I'm thinking of people like Emma Mulqueeny and Chris Thorpe who have blogged on the subject of what not to do (eg https://mulqueeny.wordpress.com/2010/11/18/developers/).

“The event may last a single day (e.g., 12), an entire week[9], or longer[18]. The number of participants may range from a few dozen (e.g., 8, 27)” - Should those references be linked?

“So, if we look at hackathons as proposal germinators, this is rather high impact. The total award amount for the two National Science Foundation grants is approximately $7.5 M. By comparison, the total amount spent on the nine NESCent hackathons was roughly $250,000.” - I wonder what the fair comparator would be here? Is ROI a sensible measure at all. It seems to me that subsequent (grant) income a good proxy for activity which is something you might want to measure.

Is the work clearly and accurately presented and does it cite the current literature?
Partly

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Reviewer Report 14 June 2017

https://doi.org/10.5256/f1000research.12340.r23270

© 2017 Amsen E. This is an open access peer review report distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Eva Amsen
The Institute of Cancer Research, London, UK

This article describes a post-hoc analysis of nine hackathons organised by NESCent between 2006 and 2015 to work on the development of software tools for evolutionary biology. Although it's presented as a research article, the main value of this article is the "Lessons Learned" section, which, in addition to the "Concise Guide" available on Github, serves as useful guidelines for anyone planning to organise hackathon-style events for academics. I particularly like the detailed description of the authors' approach to ensure participant engagement and to increase attendance by underrepresented groups, as this is something many event organisers can benefit from.

The authors address the shortcomings of the study by pointing out that an objective systematic analysis of these nine events (which they organised themselves over a period of several years) was not possible. However, thanks to its detailed descriptions and evaluations, this article and the accompanying guidelines may help others perform a more systematic analysis of future hackathons.

My only comments (below) are suggestions to make the text easier to read. In particular, the Introduction could be more informative to readers who are new to hackathons, by including some information that in the present version is mentioned much later in the article:

1. Who was the intended audience for the NESCent hackathons? The introduction mentions "developers and end-users", but it is not until Lesson 4 that the reader finds out they were "a mix of faculty, postdocs, students and research staff". Considering the academic audience of the article, this might be useful to mention earlier.

2. Another point to potentially address in the article is the differences between a hackathon and any other type of scientific workshop. For example, the participant-driven nature of hackathons (first mentioned in Lesson 6) is key to understanding the concept of the event. If this is something that needs to be explained to new hackathon participants, perhaps it also needs to be explained to new readers.

Suggested minor cosmetic changes:

- Some in-text citations in the introduction are not linked to their respective references. There

are three of these close together, starting with “The event may last a single day (e.g. 12)....”

- There is a typo in reference 6: First author should be “McArthur”, not “McArthu”
- It takes some effort to find the sample advertisements mentioned in the “recruiting participants” section. This could be improved with a direct link to the corresponding files.

Is the work clearly and accurately presented and does it cite the current literature?
Yes

Is the study design appropriate and is the work technically sound?
Yes

Are sufficient details of methods and analysis provided to allow replication by others?
Yes

If applicable, is the statistical analysis and its interpretation appropriate?
Not applicable

Are all the source data underlying the results available to ensure full reproducibility?
Yes

Are the conclusions drawn adequately supported by the results?
Yes

Competing Interests: I previously worked for F1000 and F1000Research in outreach and community management roles, but I have not been employed by the company since January 2016, and have no current insight or stakes in the referee process of F1000Research articles.

Reviewer Expertise: Science communication, with expertise in organising participant-driven events for academics

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

The benefits of publishing with F1000Research:

- Your article is published within days, with no editorial bias
- You can publish traditional articles, null/negative results, case reports, data notes and more
- The peer review process is transparent and collaborative
- Your article is indexed in PubMed after passing peer review
- Dedicated customer support at every stage

For pre-submission enquiries, contact research@f1000.com

F1000Research