Collaborative community coding events in open source biological research

Brad Chapman
Harvard Chan School Bioinformatics Core
http://bioinformatics.sph.harvard.edu/
bchapman@hsph.harvard.edu

Abstract

The Open Bioinformatics Foundation Codefest is a multiple day collaborative working session. Codefest provides a venue for real time collaboration between researchers who have established relationships through decentralized open source work, as well as a place for new developers to integrate with a welcoming community. We'll describe the unique collaborative structure of Codefest and discuss approaches to help improve the event to sustain long term collaborations while training a diverse set of attendees.

Author Keywords

bioinformatics, collaboration, training, diversity, open source, biology

Background and motivation

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 event called the OpenBio Codefest (https://www.open-bio.org/wiki/Codefest). This event contin-

ued the past 8 years in a wide diversity of locations, with the most recent taking place at a non-profit, community-run hackerspace in Prague (https://www.open-bio.org/wiki/Codefest_2017).

This summer, we've combined with another open source community to create a full bioinformatics community conference including dedicated training, traditional conference talks and four days of collaboration (https://gccbosc2018.sched.com/).

Our goals at the CHI 2018 Hackathon Workshop are to describe the unique collaborative structure of Codefest, connect with other organizers building long term community relationships through collaborative events, and learn about how we can improve at training a diverse set of attendees.

Collaborative event design

Codefest initially started as a space for community members who were already collaborating remotely to sit together and work. Over time, it expanded to better incorporate new members into the community by serving as a fun and open environment for sharing work and meeting like-minded researchers.

We plan to share some unique design elements we've learned in organizing Codefest:

- 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. These standards have been essential for forming new long term collaborations for building necessary research infrastructure.

Training and community building

The biological problems we work on at Codefest require collaboration across a diverse set of research areas. We're continually focused on strengthening and improving our community and are hoping to learn from other organizers at the Workshop:

- 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 underrepresented 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. At this year's upcoming Codefest we plan to evaluate the impact of having dedicated training prior to the collaborative hands on event.

- 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 share 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.