

## The Bioinformatics Open Source Conference (BOSC) 2013

Nomi L. Harris<sup>1,\*</sup>, Peter J. A. Cock<sup>2</sup>, Brad A. Chapman<sup>3</sup>, Jeremy Goecks<sup>4</sup>,  
Hans-Rudolf Hotz<sup>5</sup> and Hilmar Lapp<sup>6</sup>

<sup>1</sup>Physical Biosciences Division, Lawrence Berkeley National Laboratory, MS 972R525, Berkeley, CA 94720, USA,

<sup>2</sup>Information and Computational Sciences, The James Hutton Institute, Dundee DD2 5DA, UK, <sup>3</sup>Bioinformatics Core, Harvard School of Public Health, Boston, MA 02115, USA, <sup>4</sup>The Computational Biology Institute, George Washington University, Washington, DC 20052, USA, <sup>5</sup>Bioinformatics Department, Friedrich Miescher Institute for Biomedical Research, Maulbeerstrasse 66, 4058 Basel, Switzerland and <sup>6</sup>National Evolutionary Synthesis Center (NESCent), Durham, NC 27705, USA

**Contact:** nlharris@lbl.gov

In July 2013, more than 100 bioinformatics researchers, developers and users of Open Source Software gathered in Berlin, Germany, to attend the 14th Annual Bioinformatics Open Source Conference (BOSC, [http://www.open-bio.org/wiki/BOSC\\_2013](http://www.open-bio.org/wiki/BOSC_2013), Harris *et al.*, 2013). Since its inception in 2000, BOSC has provided bioinformatics developers with a forum for communicating the results of their latest efforts to the wider research community, and a focused environment for developers and users to interact and share ideas about standards, software development practices and practical techniques for solving bioinformatics problems. BOSC includes a 2 day ‘CodeFest’ preceding the formal conference, which provides a venue for developers to meet in person to work on or plan joint projects (Möller *et al.*, 2013).

The session topics at BOSC 2013 included Cloud and Parallel Computing, Visualization, Software Interoperability, Genome-scale Data Management, a session for updates on ongoing Open Source projects and two new sessions: Open Science and Reproducible Research and Translational Bioinformatics, recognizing the growing use of computational biology in medical applications. Slides from all of the presentations are available on the BOSC Web site, along with some of the posters and video recordings for selected talks, including the keynotes and panel discussion.

Open Science is a movement dedicated to making all aspects of scientific knowledge production freely available for reuse and extension, including scientific data, methods and analyses. In response to the increasing traction that this movement has gained, BOSC 2013 featured a new session devoted explicitly to Open Science. One of the objectives of Open Science is the wider issue of making published research reproducible. Aside from openness in software licensing, this also includes openness of data and unhindered access to scientific papers themselves (Open Access). When researchers can freely access publications and the source code and data that support them, it becomes possible for them to recreate the steps that the authors went

through to reach their conclusions and to then go beyond them. In this way, Open Science not only stands to provide the value of validating published results by recreating them but also to accelerate the pace of scientific discovery, by enabling researchers to more effectively build on the results of previous work, rather than having to reinvent tools and reassemble datasets.

Each day of BOSC traditionally starts off with a keynote talk by a person of influence in Open Source bioinformatics. BOSC 2013’s first keynote speaker was Cameron Neylon, the Advocacy Director for the Public Library of Science, who is a prominent advocate for open science. Neylon discussed the cultural issues that are hindering open science, and spoke about the potential of openness in scientific collaborations for generating impact. Our second keynote speaker was Sean Eddy, a group leader at the Howard Hughes Medical Institute’s Janelia Farm, who is perhaps best known as the author of the HMMER software suite (Eddy, 2011). He discussed how his own experience and practices show that dedicating effort to thorough engineering in tool development—which is often shunned as uninteresting and unpublishable—can be a key to creating a lasting impact.

To stimulate discussion on controversial or multifaceted topics, BOSC includes a panel, in which panelists representing a range of viewpoints, answer questions from the audience. This year’s panel was on Strategies for Funding and Maintaining Open Source Software. To secure continued funding for a software project, researchers must be able to demonstrate its impact. The panelists explored ways to quantify usage of one’s software as a measure of impact. They agreed that traditional publications, and tracking their citations, still play an important role in publicizing and demonstrating the use of one’s software, but they are not the only metric.

With the increasing reliance of more and more fields of biology on computational tools to manage and analyze their data, BOSC seems assured to stay relevant to life science, and thus to life scientists.

### Additional Information

Plan to join us for BOSC: 15th Annual Bioinformatics Open Source Conference in Boston, USA, July 11–12, 2014. For

\*To whom correspondence should be addressed.

more information visit [http://www.open-bio.org/wiki/BOSC\\_2014](http://www.open-bio.org/wiki/BOSC_2014).

## ACKNOWLEDGEMENTS

The authors thank all the current and past members of the BOSC organizing and abstract review committees, the Codefest organizers and hosts and their panelists, speakers and delegates for making every BOSC a stimulating and informative meeting. They are grateful to Eagle Genomics for repeatedly sponsoring student travel awards and to ISCB and the ISMB meeting organizers—especially Steven Leard—for their part in making BOSC happen.

**Funding:** NLH was supported by the U.S. Department of Energy, Office of Biological and Environmental Research under Contract DE-AC02-06CH11357. PJC was supported by the Scottish Government Rural and Environmental Research and Analysis Directorate. BC has no external funding to

report. JG was supported in part by grant number HG005542 from the National Human Genome Research Institute, National Institutes of Health, as well as grants HG005133 and HG006620. HRH was supported by the Novartis Research Foundation HL was supported by the National Evolutionary Synthesis Center (NESCent), NSF #EF-0905606.

*Conflicts of interest:* none declared.

## REFERENCES

- Eddy,S. (2011) Accelerated Profile HMM Searches. *PLoS Comput. Biol.*, **7**, e1002195. <http://dx.doi.org/10.1371/journal.pcbi.1002195> (14 July 2014, date last accessed).
- Harris,N. et al. (2013) The Bioinformatics Open Source Conference (BOSC) 2013. *PeerJ PrePrints*, **1**, e83v1. <http://dx.doi.org/10.7287/peerj.preprints.83v1> (14 July 2014, date last accessed).
- Möller,S. et al. (2013) Sprints, Hackathons and Codefests as community gluons in computational biology. *EMBnet.J.*, **19**, 40–42. <http://journal.embnet.org/index.php/embnetjournal/article/view/726/998> (14 July 2014, date last accessed).