# 13ᵗʰ Annual Bioinformatics Open Source Conference
# BOSC 2012 Schedule

**Day 1 (Friday, July 13, 2012)**

| Time | Title | Speaker or Session Chair |
|---|---|---|
| 7:30-9:00 | **Registration** | |
| 9:00-9:15 | **Introduction and Welcome** | Nomi Harris (Chair, BOSC 2012) |
| 9:15-10:15 | **Keynote: Science Wants to Be Open - If Only We Could Get Out of Its Way** | Jonathan Eisen |
| 10:15-10:45 | *Coffee Break* | |
| 10:45-12:30 | **Session: Cloud and Parallel Computing** | Chair: Richard Holland |
| 10:45-11:10 | Cloudgene - an execution platform for MapReduce programs in public and private clouds | Sebastian Schoenherr |
| 11:10-11:35 | Data reduction and division approaches for assembling short-read data in the cloud | C. Titus Brown |
| 1135-12:00 | MGTAXA - a toolkit and a Web server for predicting taxonomy of the metagenomic sequences with Galaxy front-end and parallel computational back-end | Andrey Tovchigrechko |
| 12:00-12:15 | Workflows on the Cloud - Scaling for National Service | Katy Wolstencroft |
| 12:15-12:30 | How to use BioJava to calculate one billion protein structure alignments at the RCSB PDB website | Andreas Prlic |
| 12:30-1:30 | *Lunch* | |
| 12:30-2:00 | **Poster Session I** | |
| 2:00-3:30 | **Session: Genome-scale Data Management** | Chair: Ronald Taylor |
| 2:00-2:15 | Using HDF5 to work with large quantities of biological data | Dana Robinson |
| 2:15-2:30 | Large scale data management in Chipster 2 workflow environment | Aleksi Kallio |
| 2:30-2:45 | JBrowse 2012 | Ian Holmes |
| 2:45-3:00 | Khmer: A probabilistic approach for efficient counting of k-mers | Qingpeng Zhang |
| 3:00-3:15 | AmiGO 2: a document-oriented approach to ontology software and escaping the heartache of an SQL backend | Seth Carbon |
| 3:15-3:30 | Discovery of motif-based regulatory signatures in whole genome methylation experiments | Jens Lichtenberg |

| Time | Title | Speaker or Session Chair |
|---|---|---|
| 3:30-4:00 | *Coffee Break* | |
| 4:00-5:30 | **Session: Linked Data and Translational Knowledge Discovery** | Chair: Jan Aerts |
| 4:00-4:25 | The open source ISA metadata tracking framework: from data curation and management at the source, to the linked data universe | Philippe Rocca-Serra |
| 4:25-4:50 | Automated infrastructure for custom variant comparison and analysis | Brad Chapman |
| 4:50-5:05 | KUPKB: Sharing, Connecting and Exposing Kidney and Urinary Knowledge using RDF and OWL | Julie Klein |
| 5:05-5:20 | eagle-i: development and expansion of a scientific resource discovery network | Sophia Cheng |
| 5:20-5:30 | Report on Codefest 2012 | Brad Chapman |
| 5:30-5:40 | Introduction to the Open Bioinformatics Foundation (O|B|F) | Hilmar Lapp (President, O|B|F) |
| 5:40-6:40 | **Poster Session II** | |
| 5:40-6:40 | **BOFs** | |
| 7:00 | Pay-your-own-way BOSC dinner ($25) RSVP at http://bit.ly/BOSC2012-dinner | George's Greek Café, 135 Pine Ave. |

## Day 2 (Saturday, July 14, 2012)

| Time | Title | Speaker or Session Chair |
|---|---|---|
| 8:45-8:50 | **Announcements** | Nomi Harris |
| 8:50-9:50 | **Keynote: If I Build It Will They Come?** | Carole Goble |
| 9:50-10:15 | Pistoia Alliance Sequence Squeeze: Using a competition model to spur development of novel open-source algorithms | Richard Holland |
| 10:15-10:45 | *Coffee Break* | |
| 10:45-12:30 | **Session: Software Interoperability** | Chair: Hilmar Lapp |
| 10:45-11:00 | GenomeSpace: An open source environment for frictionless bioinformatics | Michael Reich |
| 11:00-11:15 | Galaxy Project Update | Dannon Baker |
| 11:15-11:30 | Zero to a Bioinformatics Analysis Platform in 4 Minutes | Enis Afgan |

| Time | Title | Speaker or Session Chair |
|---|---|---|
| 11:30-11:55 | PyPedia: A python crowdsourcing development environment for bioinformatics and computational biology | Alexandros Kanterakis |
| 11:55-12:10 | InterMine - Embeddable Data-Mining Components | Alexis Kalderimis |
| 12:10-12:25 | Creating biology pipelines with BioUno | Bruno Kinoshita |
| 12:30-1:30 | *Lunch* | |
| 12:30-1:45 | **Poster Session III** | |
| 1:45-3:30 | **Session: Bioinformatics Open Source Project Updates** | Chair: Jeremy Goecks |
| 1:45-2:00 | Mobyle Web Framework: New Features | Hervé Ménager |
| 2:00-2:15 | Biopython Project Update | Eric Talevich |
| 2:15-2:40 | Biogem,Ruby UCSC API, and BioRuby | Hiroyuki Mishima |
| 2:40-2:55 | A Framework for Interactive Visual Analysis of NGS Data using Galaxy | Jeremy Goecks |
| 2:55-3:10 | Why Scientists Should Contribute to Wikipedia | Spencer Bliven |
| 3:10-3:25 | scabio - a framework for bioinformatics algorithms in Scala | Markus Gumbel |
| 3:30-4:00 | *Coffee Break* | |
| 4:00-4:30 | **Session: Lightning Talks** | Chair: Kam Dahlquist |
| 4:00-4:07 | bioKepler: A Comprehensive Bioinformatics Scientific Workflow Module for Distributed Analysis of Large-Scale Biological Data | Jianwu Wang |
| 4:07-4:14 | (Last-minute lightning talk) | |
| 4:14-4:21 | (Last-minute lightning talk) | |
| 4:21-4:28 | Bioinformatics Testing Consortium: Codebase peer-review to improve robustness of bioinformatic pipelines | Ben Temperton |
| 4:30-5:30 | **Panel: Bioinformatics Paper Reviews--Openness of Standards and Standards of Openness** | *Moderator*: Brad Chapman *Panelists*: Titus Brown, Iain Hrynaszkiewicz, Hilmar Lapp, Scott Markel, Ben Temperton |
| 5:30-5:40 | Presentation of awards | Nomi Harris |
| 5:40-6:40 | **BOFs** | |

Any last-minute schedule updates will be posted at http://www.open-bio.org/wiki/BOSC_2012_Schedule

# Posters

| Number | Poster Title and Author |
| --- | --- |
| 1 | Cloudgene - an execution platform for MapReduce programs in public and private clouds (Sebastian Schoenherr) |
| 2 | Data reduction and division approaches for assembling short-read data in the cloud (C. Titus Brown) |
| 3 | Workflows on the Cloud: Scaling for National Service (Katy Wolstencroft) |
| 4 | Using HDF5 to Work With Large Quantities of Biological Data (Dana Robinson) |
| 5 | Large scale data management in Chipster 2 workflow environment (Aleksi Kallio) |
| 6 | Khmer: A probabilistic approach for efficient counting of k-mers (Qingpeng Zhang) |
| 7 | Discovery of motif-based regulatory signatures in whole genome methylation experiments (Jens Lichtenberg) |
| 8 | Pistoia Alliance Sequence Squeeze: Using a competition model to spur development of novel open-source algorithms (Richard Holland) |
| 9 | GenomeSpace: An open source environment for frictionless bioinformatics (Michael Reich) |
| 10 | Zero to a Bioinformatics Analysis Platform in 4 Minutes (Enis Afgan) |
| 11 | PyPedia: A python crowdsourcing development environment for bioinformatics and computational biology (Alexandros Kanterakis) |
| 12 | InterMine - Embeddable Data-Mining Components (Alexis Kalderimis) |
| 13 | Creating biology pipelines with BioUno (Bruno Kinoshita) |
| 14 | Mobyle Web Framework: New Features (Hervé Ménager) |
| 15 | Biogem,Ruby UCSC API, and BioRuby (Hiroyuki Mishima) |
| 16 | Why Scientists Should Contribute to Wikipedia (Spencer Bliven) |
| 17 | scabio - a framework for bioinformatics algorithms in Scala (Markus Gumbel) |
| 18 | bioKepler: A Comprehensive Bioinformatics Scientific Workflow Module for Distributed Analysis of Large-Scale Biological Data (Jianwu Wang) |
| 19 | Bioinformatics Testing Consortium: Codebase peer-review to improve robustness of bioinformatic pipelines (Ben Temperton) |
| 20-25 | *Walk-in posters* |