

Flatware

Report of Contributions

Contribution ID: 1

Type: **not specified**

SCC Tool Talk: Jupiter Widgets

Presenter: WATTERS, Aaron

Contribution ID: 2

Type: **not specified**

CCA Project Talk: Astropy

Presenters: PRICE-WHELAN, Adrian; CRUZ, Kelle

Contribution ID: 3

Type: **not specified**

Lightning Talk:CCQ

Presenter: CARLEO, Guiseppe

Contribution ID: 4

Type: **not specified**

Lightning Talk: CCB-Biophysical Modeling

Presenter: YAN, Wen

Contribution ID: 5

Type: **not specified**

CCB Project Talk: 20 Years of Biomolecular Modeling in Rosetta

Presenters: RENFREW, Doug; KOEHLER-LEMAN, Julia; MULLIGAN, Vikram

Contribution ID: 6

Type: **not specified**

Lightning Talk: CCA

Presenters: SIMON, Dylan; GENEL, Shy

Session Classification: Lightning Talks

Contribution ID: 7

Type: **not specified**

Lightning Talk: CCM?

Presenter: JUN, James

Session Classification: Lightning Talks

Contribution ID: 8

Type: **not specified**

CCQ Project Talk:

Presenter: PARCOLLET, Olivier (CCQ)

Contribution ID: 9

Type: **not specified**

Lightning Talk: CCA

Presenter: BRYANT, Greg

Contribution ID: **10**

Type: **not specified**

CCA Project TALK: Daedalus

Presenters: OISHI, Jeff; BURNS, Keaton

Contribution ID: 11

Type: **not specified**

Interaction with users within open-source software for scientific applications

Presenters: GIOVANNUCCI, Andrea; PNEVMATIKAKIS, Eftychios

Contribution ID: 12

Type: **not specified**

Lightning Talk: CCQ

Presenters: SIMON, Dylan; WENTZELL, Nils

Contribution ID: 13

Type: **not specified**

SCC Tool Talk: MyPy

Presenter: GUNN, Pat

Contribution ID: 14

Type: **not specified**

Project Talk CCB: Humanbase

Presenter: WONG, Aaron

Contribution ID: 15

Type: **not specified**

Lightning Talk: CCM?

Presenter: MAGLAND, Jeremy

Contribution ID: 16

Type: **not specified**

Project Talk CCQ: Itemsor

Presenter: STOUDENMIRE, Miles

Contribution ID: 17

Type: **not specified**

Lightning Talk: CCB-Biophysical

Presenter: STEIN , David

Flatware

/ Report of Contributions

Speaker

Contribution ID: **18**

Type: **not specified**

Speaker

Presenters: CHEN, Kathleen; CARRIERO, Nick (FI SCC)

Contribution ID: 19

Type: **not specified**

CCB Genomics

Primary author: CHEN, Kathleen

Presenter: CHEN, Kathleen

Contribution ID: 20

Type: **not specified**

CCQ NetKet Software

Primary author: CARLEO, Guiseppe

Presenter: CARLEO, Guiseppe

Contribution ID: 21

Type: **not specified**

CCB Biophysical Modeling

Primary author: YAN, Wen

Presenter: YAN, Wen

Contribution ID: 22

Type: **not specified**

CCA Lightning Talk

Contribution ID: 23

Type: **not specified**

CCM

Contribution ID: 24

Type: **not specified**

Dedalus: A flexible framework for solving differential equations using spectral methods

Wednesday, October 24, 2018 9:00 AM (1h 30m)

Dedalus is an open-source library for solving partial differential equations using global spectral methods. These methods are well-suited to solving smooth PDEs, such as those describing fluid flows at low Mach-numbers, with very high accuracy. Dedalus is written in Python for ease-of-use and wraps C libraries such as FFTW and MPI for performance on large-scale HPC systems. The code has been used to study problems in a number of fields including astrophysics, oceanography, atmospheric science, biological fluid dynamics, and plasma physics. We plan to discuss our experiences developing Dedalus as an open-source tool and several issues we are currently facing, including:

- Designing for a balance between capability and maintainability
- Generalized equations and timestepping routines
- Automatic MPI parallelization
- Preventing “feature-creep”
- Using a high-level language for high-performance applications
- Python optimization
- Typed-Python via Cython
- Wrapping C libraries
- Balancing ease-of-installation with dependency-optimization
- Achieving high performance on laptops, desktops, and clusters
- Docker & conda distribution
- Supporting a user-base that is substantially larger than the developer-base
- Encouraging public posts to user groups / issue trackers
- Time spent on user support
- Dealing with citations, authorship, etc.

Primary authors: OISHI, Jeff; BURNS, Keaton

Presenters: OISHI, Jeff; BURNS, Keaton

Session Classification: Projects

Contribution ID: 25

Type: **not specified**

20 Years of Macromolecular Modeling in Rosetta

Wednesday, October 24, 2018 1:00 PM (1h 30m)

Rosetta is one of the largest software suites for macromolecular modeling with 3 million lines of code and many state-of-the-art protocols. It is developed by the RosettaCommons, a community of developers from 60 laboratories worldwide. Since the mid 1990's, Rosetta has been primarily developed in an academic environment by scientists with backgrounds in chemistry, biology, physiology, physics, engineering, mathematics, computer science and related disciplines. Challenges in scientific software development are many developers' lack of formal training in software engineering or computer science and the academic environments' under-appreciation of sustainability and maintainability of tools developed for basic science research.

Here we present lessons learned from a period of over two decades in how to develop advanced scientific software in a global community with hundreds of developers. We address aspects like version control, licensing, testing, documentation, maintenance and a variety of community features such as conferences, training, hackathons, user interaction, as well as outreach and diversity efforts.

Primary authors: RENFREW, Doug; KOEHLER LEMAN, Julia; MULLIGAN, Vikram

Presenters: RENFREW, Doug; KOEHLER LEMAN, Julia; MULLIGAN, Vikram

Session Classification: Projects

Contribution ID: 26

Type: **not specified**

Jupyter Widgets

Wednesday, October 24, 2018 3:00 PM (1 hour)

This talk will discuss Jupyter widgets

- What is Jupyter and how do widgets fit into Jupyter?
- Some examples of different kinds of widgets.
- Example user stories for using widgets.
- Simple examples of creating and composing widgets.
- How to create interactive diagrams with widgets.
- How to encapsulate external Javascript Libraries in widgets.

Primary author: WATTERS, Aaron

Presenter: WATTERS, Aaron

Session Classification: Tool Talks

Contribution ID: 27

Type: **not specified**

TRIQS : a Toolbox for Research in Interacting Quantum Systems.

Thursday, October 25, 2018 9:00 AM (1h 30m)

After a general introduction to the challenges of the quantum many-body problem, and to some directions of research developed at CCQ, I will present the TRIQS project. On the technical side, I will discuss several general topics, including Python/C++ communication, hdf5, modern C++ and the associated tools to increase code quality.

Primary author: PARCOLLET, Olivier (CCQ)

Presenter: PARCOLLET, Olivier (CCQ)

Session Classification: Projects

Contribution ID: 28

Type: **not specified**

CCA

Contribution ID: 29

Type: **not specified**

CCB Neuroscience

Contribution ID: **30**

Type: **not specified**

TDB

Contribution ID: 31

Type: **not specified**

The Astropy Project: a community effort to develop a common core package for Astronomy

Thursday, October 25, 2018 2:00 PM (1h 30m)

The Astropy Project is a community effort to develop an interoperable ecosystem of open-source tools to enable astronomical research and education from the Python programming language. At its core is the Astropy package, a Python package that provides much of the base functionality needed by researchers and developers of more specialized packages. The Astropy project and package have now formally existed for nearly 6 years, and the community of contributors and users has changed significantly during this time. We will discuss the initial need for the project, the initial development, and recent efforts towards making the existing tools more stable and supported. We will also discuss the development of the community and the future sustainability of this and other core open-source projects developed by active researchers.

Primary authors: PRICE-WHELAN, Adrian; CRUZ, Kelle

Presenters: PRICE-WHELAN, Adrian; CRUZ, Kelle

Session Classification: Projects

Flatware

/ Report of Contributions

TBD

Contribution ID: **32**

Type: **not specified**

TBD

Contribution ID: 33

Type: **not specified**

CCQ Singularity Indeed

Primary author: WENTZELL , Nils

Presenter: WENTZELL , Nils

Contribution ID: 34

Type: **not specified**

HumanBase: A portal for data-driven predictions of gene function, regulation, and interactions

Friday, October 26, 2018 9:00 AM (1h 30m)

HumanBase (hb.flatironinstitute.org) is a comprehensive resource for biomedical researchers interested in exploring expression, function, regulation and interactions of human genes, particularly in the context of specific tissues/cell-types and human disease. Data-driven integrative analyses underlying HumanBase are especially powerful because they reach beyond existing “biological knowledge” represented in the literature to identify novel associations that are not biased toward well-studied areas of biomedical research. HumanBase integrates data from more than 38,000 genomic experiments and more than 14,000 scientific publications to uncover genes’ tissue-specific function and roles in disease, inter-relationships between genes, and the gene expression effects of genetic variants. We will discuss how HumanBase addresses an unmet need among biologists, the development of the web-based system, and the challenges and pitfalls of developing this public resource.

Primary author: WONG, Aaron

Presenters: WONG, Aaron; FUNK, Julien

Session Classification: Projects

Contribution ID: 35

Type: **not specified**

CCQ

Contribution ID: 36

Type: **not specified**

CCM

Contribution ID: 37

Type: **not specified**

ITensor

Friday, October 26, 2018 1:30 PM (1h 30m)

ITensor is a library for creating high-performance codes for tensor network algorithms. ITensor facilitates rapid prototyping as well as long-term maintainability. After reviewing many-body quantum physics from a tensor mathematics perspective, I will introduce ITensor and what distinguishes it from most other tensor libraries. Then I will discuss good and less good design decisions made through the course of ITensor's development and draw some general lessons. These decisions touch on areas such as choice of language, interface design, and documentation. To conclude, I will highlight future goals and topics where we could benefit from the expertise of others at Flatiron.

Presenter: STOUDENMIRE, Miles**Session Classification:** Projects

Contribution ID: **38**

Type: **not specified**

CCB-Biophysical Modeling

Contribution ID: 39

Type: **not specified**

IronClust: Real-time, drift-resistant spike-sorting based on density-based clustering

Software engineering focus

Robust performance under physical drift over time

Validation using simulated and biophysical ground-truth datasets

Realtime performance up to 1000 channels using parallel computing hardware (GPU, CPU)

Presenter: JUN, James

Contribution ID: 40

Type: **not specified**

Selene: A framework for training sequence level deep learning networks

Wednesday, October 24, 2018 11:00 AM (15 minutes)

To enable the application of deep learning in biology, we developed Selene, a PyTorch-based library for fast and easy development, training, and application of deep learning model architectures to sequence-level (e.g. DNA, RNA) datasets. In this presentation, I will discuss how we designed Selene to support sequence-based deep learning across a broad range of biological questions and made the library accessible to users at different levels of computational proficiency.

Presenter: CHEN, Kathleen

Session Classification: Lightning Talks

Contribution ID: 41

Type: **not specified**

NetKet: challenges and desiderata in designing machine learning software for quantum physics

Thursday, October 25, 2018 4:15 PM (15 minutes)

I will discuss the open-source project NetKet (<https://www.netket.org/>), and the challenges in designing a software flexible enough for research purposes while keeping the pace of large-scale machine learning software developed by industry. I will also discuss the strategies that we have put in place to stimulate external collaborations to the project.

Presenter: CARLEO, Giuseppe

Session Classification: Lightning Talks

Contribution ID: 42

Type: **not specified**

SafeFFT: A thread safe interface to FFTW and MKL

Wednesday, October 24, 2018 11:30 AM (15 minutes)

Running FFTs, possibly with different sizes and types, from multiple threads is becoming necessary to improve both the program flexibility and execution efficiency, with an increasing number of CPU cores on a typical SMP machine. However, the native interface in FFTW and Intel MKL does not fully support this use scenario due to thread safety issues. To address this issue, SafeFFT is designed as a simple C++ interface that allows various FFTs can be planned and executed from multiple OpenMP threads safely. The internal implementation is based on a hash map with a custom OpenMP read-write lock that allows various FFT plans can be safely allocated, saved, and reused. Nested OpenMP threading is also supported to improve the utilization of CPU cores.

Presenter: YAN, Wen

Session Classification: Lightning Talks

Contribution ID: 43

Type: **not specified**

CCA Simulation Repository: common frameworks for sharing, distributing, and processing astrophysical simulation data

Thursday, October 25, 2018 4:30 PM (15 minutes)

As envisioned by the galaxy formation group, our goal is to provide tools to discover, share, and access cosmological and other astrophysical simulations through web-based platforms, enabling efficient analysis of large datasets by distributed users. Currently in the planning and prototyping phase, this will ultimately involve creating ways to query object catalogs, access portions of raw simulation volumes, and drive synthetic observation pipelines in generic ways across a variety of datasets.

Presenters: SIMON, Dylan; GENEL, Shy

Session Classification: Lightning Talks

Contribution ID: 44

Type: **not specified**

IronClust: Real-time, drift-resistant spike-sorting based on density-based clustering

Thursday, October 25, 2018 11:45 AM (15 minutes)

Software engineering focus:

Robust performance under physical drift over time

Validation using simulated and biophysical ground-truth datasets

Realtime performance up to 1000 channels using parallel computing hardware (GPU, CPU)

Presenter: JUN, James

Session Classification: Lightning Talks

Contribution ID: 45

Type: **not specified**

STARRY: Light curve modeling suite

Thursday, October 25, 2018 11:15 AM (15 minutes)

I will discuss the ongoing development of the open-source STARRY code for computation of fast light curves for stellar and exoplanet science. I will focus mainly on computational aspects, including (1) obtaining model derivatives via autodifferentiation for use in gradient-based inference schemes such as Hamiltonian Monte Carlo, (2) developing simultaneous C++, Julia, and Python interfaces to the code, and (3) fully integrating code development with paper writing using SymPy, IPython Notebooks, and Continuous Integration (CI) tools.

Presenter: LUGER, Rodrigo

Session Classification: Lightning Talks

Contribution ID: 46

Type: **not specified**

CaImAn: Large scale brain imaging data analysis for the 99%

Thursday, October 25, 2018 11:30 AM (15 minutes)

CaImAn is an open-source software framework for the analysis of brain imaging data. We will share our experience interacting with the CaImAn user-base, mostly composed of neuroscience researchers. Properly managing the interaction with users is essential to the success of scientific software, because of the received feedback, the potential contributions to the code-base, and the community perception. However, user proficiency in coding and computing is highly variable, and therefore targeted tools and communication channels should be used to deal with specific situations. After briefly introducing instruments like git issues, git wiki pages, gitter chat rooms, and git pull requests, we will present an inventory of exemplar and notable cases of user interaction.

Presenters: GIOVANNUCCI, Andrea; PNEVMATIKAKIS, Eftychios

Session Classification: Lightning Talks

Flatware

/ Report of Contributions

TBD

Contribution ID: 47

Type: **not specified**

TBD

Session Classification: Lightning Talks

Contribution ID: 48

Type: **not specified**

Containerization in Modern Scientific Applications

Wednesday, October 24, 2018 11:15 AM (15 minutes)

Scientific algorithms for the solution of the correlated many-body problem and beyond are rapidly growing in complexity. This development shows also in their respective numerical implementations, leading to growing number of library dependencies, toolchain dependencies as well as interdependencies with other scientific applications. This often makes a proper setup on workstations, but in particular on high-performance cluster machines an ambitious and sometimes insurmountable task. In this presentation I will give an introduction into modern containerization tools, in particular Docker and Singularity, that are targeted to overcome these issues.

Presenter: WENTZELL , Nils

Session Classification: Lightning Talks

Contribution ID: 49

Type: **not specified**

TBD

Session Classification: Lightning Talks

Contribution ID: 50

Type: **not specified**

Auxiliary Field Quantum Monte Carlo software

Friday, October 26, 2018 11:15 AM (15 minutes)

I will present a quantum Monte Carlo software for the strongly-correlated many-electron systems. The software focuses on ab initio simulations for realistic materials and quantum chemistry systems. I will also talk about massively parallel design in the software.

Presenter: SHI, Hao

Session Classification: Lightning Talks

Contribution ID: 51

Type: **not specified**

A spike sorting meta package and website for algorithm comparison

Friday, October 26, 2018 11:30 AM (15 minutes)

Spike sorting is an crucial component of most neurophysiology pipelines that precedes downstream analysis of neural firing data. With a dozen or so spike sorting software packages in the mix, there is little to no consensus on which algorithm is most suitable, depending on the experimental setup. This is due to a number of factors including lack of realistic ground truth recordings, no clear consensus on file formats, software installation challenges, and little consensus on evaluation metrics. While we are developing and maintaining two algorithms in house (MountainSort and IronClust), we are also working on a meta package that includes all automated spike sorting algorithms wrapped in single python package with common tools for visualization and file I/O. We plan to host a website providing a rich, interactive comparison of these algorithms applied to standard ground truth datasets (both synthetic and real).

Presenter: MAGLAND, Jeremy

Session Classification: Lightning Talks

Contribution ID: 52

Type: **not specified**

Welcome

Wednesday, October 24, 2018 8:50 AM (10 minutes)

Presenter: CARRIERO, Nick (FI SCC)

Session Classification: Welcome and Opening Remarks

Contribution ID: 53

Type: **not specified**

A flexible and fast package for boundary integral equations in two dimensions

Friday, October 26, 2018 11:45 AM (15 minutes)

The solution of certain elliptic partial differential equations (Laplace, Stokes, Helmholtz, etc.) provides one of the primary building blocks necessary for the study of a wide range of physical problems. For simple domains, the solution of these equations is trivial. On complex domains, this is not the case, and many researchers have depended on methods that are slow, inaccurate, or both. This library aims to provide a family of simple to use routines which enable non-experts to use fast and accurate algorithms for solving these kinds of problems in two spatial dimensions.

Presenter: STEIN, David

Session Classification: Lightning Talks

Contribution ID: 54

Type: **not specified**

MyPy

Thursday, October 25, 2018 5:00 PM (1 hour)

Presenter: GUNN, Pat

Session Classification: Tool Talks