

Paris-Saclay - Center for Data Science

Journée sur le thème des outils

26 Octobre 2015

LAL Auditorium, Université Paris-Sud

Keynote Speakers



Fernando Pérez is a staff scientist at Lawrence Berkeley National Laboratory and a founding investigator of Berkeley Institute for Data Science (BIDS). After a PhD in Physics from the University of Colorado, Boulder, during his postdoctoral research at the CU Applied Mathematics Department, he developed novel algorithms for the fast and accurate solution of partial differential equations. Before coming to BIDS, he was a researcher at the Helen Wills Neuroscience Institute, working on problems at the intersection of applied mathematics, scientific computing, and the analysis of brain imaging data.

Today, his research focuses on creating tools for modern computational research and data science across domain disciplines, with an emphasis on high-level languages, literate computing, and reproducible research. He created IPython when he was graduate student in 2001 and continues to lead it as it evolves into the Jupyter Project, now as an open, collaborative effort with a talented team that does all the hard work. He regularly lectures about scientific computing and data science and is a member of the Python Software Foundation as well as a founding member of the Numfocus Foundation. He is the recipient of the 2012 Award for the Advancement of Free Software from the Free Software Foundation.



Andreas Müller is a Research Engineer at the NYU Center for Data Science, building a group to work on open source software for data science. Previously he was a Machine Learning Scientist at Amazon, working on computer vision and forecasting problems. He is one of the core developers of the scikit-learn machine learning library, and have maintained it for several years.

His mission is to create open tools to lower the barrier of entry for machine learning applications, promote reproducible science and democratize the access to high-quality machine learning algorithms.

Program

8h30 - 9h00 Welcome participants

9h00 - 9h15 Alexandre Gramfort, Welcome talk

9h15 - 10h15 Fernando Perez

10h15 - 10h30 Oscar Najera: *Sphinx-Gallery: Package documentation made easy*

10h30 - 11h00 Coffee Break

11h00 - 11h15 Mehdi Cherti: *Py-Earth: Multivariate Adaptive Regression Splines in Python*

11h15 - 11h30 Lorenzo Desantis: *MNE from shell scripts and Unix commands to Python*

11h30 - 12h00 Sarah Cohen Boulakia: *Reproducible research in Bioinformatics: Scientific Workflows, Provenance and beyond*

12h00 - 12h15 Karin Dassas: *LoOPS Network for developers at Paris-Saclay*

12h15 - 12h30 Cécile Germain: The io.datascience DaaS platform

12h30 - 14h00 Lunch break at Proto 204

14h00 - 14h20 Miguel Colom, *CMLA: Building a service-oriented platform for online physiological data analysis*

14h20 - 14h40 Diem BUI THI, *Automation for chemical data analysis techniques*

14h40 - 15h00 Estelle Chaix, *Information Extraction Challenge for Gene Regulation Network in plant*

15h00 - 15h30 Coffee Break

15h30 - 15h50 Loic Estève, *Mining brain imaging data: lessons learned from nilearn and joblib*

15h50 - 16h10 Jonathan Duperrier & Domenico Guarino, *A customizable framework for neurophysiology data management and provenance tracking*

16h10 - 16h25 Nicolas Goix: *Anomaly detection algorithms in Scikit-Learn*

16h25 - 16h40 Romain Brault: *Structured Prediction with Opera-Lib*

16h40 - 17h30 Andreas Mueller, *Software and engineering efforts at NYU Center for Datascience*