



Laboratório de Instrumentação e Física Experimental de Partículas

# Competence Center - Big Data

---

3rd Informal Meeting - 27th October 2017

## ATTENDEES (via Vidyo)

Alexandre Lindote, Ana Peixoto, Ana Sofia Nunes, António Esteves, Celso Franco, David Fernandes, Filipe Veloso, Giles Strong, Helmut Wolters, Juanpe Araque, Lorenzo Cazon, Mário David, Nuno Castro, Orlando Oliveira, Ruben Conceição, Tiago Vale, Vladimir Solovov

## NOTES

- **Agenda:** <https://indico.lip.pt/indico/conferenceDisplay.py?confId=316>
- **Introduction:**
  - Evaluation of LIP. We should discuss the contributions and plans of our competence center.
  - It is important to discuss how to establish some dedicated computational resources at LIP for machine learning (testing, tutorials, etc).
  - Lorenzo reported that LIP will hold the week of 12th to 16th of March 2018 the event *Data Science in (Astro)Particle Physics and the Bridge to Industry*. The first three days there will be an IDPASC school in Data Science, the topics and lectures are being planned. In the last two days (15th and 16th) there will be a workshop with talks by companies working on Data Science at large (from data analysis to machine learning, statistical inference, etc). There will be also LIP-talks from the point of view of Data Science covering our main areas of research. The goal of this event is to show the synergies between fundamental research and applied Data Science in the society. A few companies have already showed interest in participating, but we'd like to call all LIP community to contribute with new ideas and key contact persons of companies that would be potentially interesting to invite, and also former Particle and Astroparticle physicists that have continued their career bringing their experience on Data Science into the industry.

- **Mário David: docker and udocker tutorial**

- Docker: light virtualization technology which became available a few years ago.
- Very easy to setup so the goal of the tutorial is to illustrate it.
- Github: public repositories.
- Tutorial available in github under LIP-Computing.
- Installing docker under different OS should be reasonably straightforward.
- Example: docker-keras-theano.
- Dockerhub can be used to create docker images.
- Docker is isolated from your host machine.
- udocker is a tool developed and maintained by Jorge Gomes.
  - Can run as user without admin privileges.
  - udocker.py used to install (can be done in user space).
  - Pull the docker image from dockerhub.
  - You can submit your jobs to any batch system using udocker and docker images.
  - Running in GPUs: NVIDIA has nvidia-docker, which is being tested.
- The computing group is available to provide support to the use of docker and udocker so please get in touch when needed.

## NEXT MEETING

Next meeting will be organized on **15th December 2017**. For this meeting the journal club will be organized by André David and it will cover the use of multivariate techniques in the H  $\rightarrow$  gamma gamma analysis in CMS. Please volunteer for the journal club during the next sessions.