

# The Fifth International School on Open Science Cloud (SOSC 2023)

The theme of the 5th edition of the international School on Open Science Cloud (SOSC) is “**Computing Models for Scientific Experiments**”. The aim of the School is to provide theoretical and practical training for students who are looking to develop their knowledge and skills on how to exploit data and computing resources for scientific applications, and how to use cloud-based infrastructures to support scientific workflows.

The 5th edition of the international School on Open Science Cloud (SOSC 2023) will be held in **Perugia, from 23 to 27 October 2023**. The school is organized by INFN, the Department of Physics and Astronomy "Augusto Righi" of the University of Bologna and the Departments of Physics and Geology of the University of Perugia.

The School is multi-disciplinary and targeted at postgraduate researchers holding a bachelor degree or equivalent in fields such as physics, statistics, computer science, computer vision, biology, medicine, bioinformatics, engineering, working at research or academic institutes, with some experience and interest in data analysis, computing or related fields. Applications by undergraduate university students will be considered depending on availability and must be accompanied by a letter of reference from a university professor. We welcome applications from all nationalities, and encourage all qualified persons to apply.

## Academic Program

The School programme, structured in both lectures and hands-on sessions, is organized over three tracks: 1) Fundamentals of software management and data handling, 2) Machine Learning methods and 3) Workflow Management in a distributed environment.

The School will start with introductory lectures on the basic concepts behind a computing model for modern scientific experiments. These will be followed by an overview of software management topics, ranging from **code portability and containers-based customization of applications** to container orchestration. Afterwards, storage solutions and best practices will be discussed in order to provide practical examples on **how to read data from a remote data service and store selected information**. **Cloud storage (such as S3)** will be considered during the hands-on sessions.

The programme will continue with a comprehensive overview of the **main Machine Learning / Deep Learning (ML/DL) models**, with a focus on basic concepts in machine learning, techniques to measure the performance of the algorithms, key ideas in the architecture of neural networks, design and implementation of models with widely-used toolkits and frameworks, basic skills to develop, deploy and maintain machine learning models in production. **Hands-on will mainly focus on Deep-Learning**. During the hands-on sessions, students will practice with various datasets spanning from [MNIST](#) to realistic physics-experiment data.

Finally, the third track will provide students with knowledge on how to organize workflows and pipelines in a distributed environment, with the objective to address the main challenges to scaling out workflow processing. To this end, an overview of distributed computing frameworks will be given and then the focus will be on **solutions based on Dask**. In order to implement a pipeline **MLFlow will be shown and used during the related hands-on session**.

The School will also host a special event reserved to the School participants, based on theoretical discussions and live demos, on novel approaches to use **FPGA** (Field Programmable Gate Array) as Machine Learning accelerators.

Alongside the SOSC tradition, the School will foresee the development of personal projects. To this extent, suitable data sets will be provided, as well as tracks to address exploratory data analysis, machine learning and workload parallelization and distribution.

#### **Prerequisites:**

- **for the hands-on sessions knowledge of Python, NumPy, Pandas and Matplotlib is expected. Familiarity with commonly used frameworks and libraries for ML/DL such as scikit-learn, Tensorflow is useful, but not mandatory. A plus would be the knowledge of docker, docker-compose, Kubernetes and cloud storage, such as S3.**

## SOSC Certificate and Diploma

An attendance certificate will be provided to the students attending the whole school programme.

The school will also issue an official "SOSC School of Computing Diploma" upon successful completion of the individual projects that will be assigned during the school.

## Fees and Enrollment

The registration fee includes morning & afternoon coffee breaks, the social dinner and participation in an evening Focus talk on FPGA, also with dinner. The **final application deadline for registrations is October 18, 2023**. Lunch, accommodation and transfer to/from the school location is not included in the fee.

The registration fee structure is as follows:

- **Early bird** (until September 15): **€200** -
- **Full ticket** (until October 6): **€300** -

Details about Registration and Payment Method can be found [here](#)

Support for a limited number of applications is possible and can be requested directly to the School secretariat.

**If you have any questions related to the registration or the payment please contact [sosc23-pc@lists.infn.it](mailto:sosc23-pc@lists.infn.it)**

## Additional Information

Participation is limited to 30 participants. In case there are more than 30 registration requests, the organizing committee will make a selection based on the scientific CV of the applicants, informing the applicants of the results.

## Venue

**INFN Perugia**, Via Alessandro Pascoli, 23c 06123 Perugia

Detailed information about the school, transportations, logistics and lectures are available at the following URL: [Travel & Accomodation](#)

## School Contact

School On Open Science Cloud 2023 - October 23th - 27th 2022

SOSC23 URL: <https://web.infn.it/SOSC23>

### **The organizing committee**

Daniele Spiga  
Davide Salomoni  
Lucio Anderlini  
Mirko Mariotti  
Daniele Bonacorsi  
Livio Fanò  
Luca Scrucca