Mosaic: Principled Foundation Models for the Physical Sciences
Principal investigator: Prof. Gilles Louppe, Université de Liège (Belgium).

I. Abstract

In an era of data-intensive science, where the volume and diversity of data is growing steadily, the obstacle to scientific progress lies not in collecting data, but in extracting coherent knowledge from it. To overcome this challenge, the Mosaic project aims to develop a rigorous methodology for **building Foundation Models (FMs) tailored to the physical sciences, with a particular focus on weather science** where the complexity and scale of data make them ideal testbeds for our research.

FMs have recently demonstrated remarkable abilities to learn general-purpose representations from large and heterogeneous sources of data, with early prototypes in the physical sciences showing promising results. Most FMs tried so far in scientific domains, however, are brute-force attempts shoehorning scientific data into deep neural networks originally designed for language and vision. Their lack of consideration for the physical laws governing natural phenomena, the instruments that produced the data, or the uncertainties inherent in a scientific analysis limits their deployment and reliability in scientific tasks. The Mosaic project aims to address these limitations by developing a rigorous methodology for building FMs aimed to represent natural phenomena, from their training on noisy, incomplete, multi-modal data to their use and certification for scientific tasks. By advancing FMs beyond brute-force approaches, we aim to bring a principled paradigm for scientific data analysis, facilitate the use of AI in science, and help reveal new scientific insights.

II. Funded positions

The Mosaic project has been funded by the F.R.S.-FNRS under an Incentive Grant for Scientific Research (MIS, Mission d'impulsion scientifique). Three fully-funded positions are available:

~~**[1] PhD position in deep learning for scientific foundation models.** We are seeking a PhD student to join the Mosaic project at the University of Liège (Belgium), focusing on developing novel representation learning approaches for scientific foundation models. The ideal candidate should have a strong background in deep learning and machine learning, with demonstrable programming skills and an interest in physical sciences. The research will involve **designing self-supervised learning algorithms for handling noisy instrumental data and developing architectures for multi-modal, multi-scale scientific data**. A significant part of the project will also involve **building foundation models from meteorological data** as a collaborative effort with the entire Mosaic team. Candidates with a background in atmospheric sciences or climate modeling are particularly encouraged to apply.~~

- ~~Starting date: as soon as possible.~~
- ~~Duration: 4 years.~~
- ~~Salary: 2500+ EUR (net) / month~~

(The PhD position is now filled!)

[2] PhD/Postdoctoral position in physics-informed foundation models. We are looking for either a PhD candidate or a postdoctoral researcher to lead research on physics-informed foundation models. The research will focus on **ensuring physical consistency in foundation models and developing rigorous certification methods for scientific applications**. The researcher will also participate in a team effort to **build foundation models for weather and climate applications**. Candidates with a background in deep learning, physics, atmospheric sciences or climate modeling are particularly encouraged to apply.

- Starting date: As soon as possible
- Duration: 4 (PhD) or 2 (postdoc) years.
- Salary: 2500+ (PhD) or 2800+ (postdoc) EUR (net) / month

[3] PhD/Postdoctoral position in foundation models for solving inverse problems. We invite applications for either a PhD candidate or a postdoctoral researcher focused on foundation models in the context of inverse problems. The research will involve **developing principled approaches for coupling foundation**

models with physical simulators and investigating their use as empirical priors for inverse problems. As part of the broader Mosaic team, the researcher will also contribute to **building foundation models for weather and climate applications**.

- Starting date: As soon as possible
- Duration: 4 (PhD) or 2 (postdoc) years.
- Salary: 2500+ (PhD) or 2800+ (postdoc) EUR (net) / month

III. Work environment

The positions are based at the University of Liège, one of Belgium's leading research universities. You will join Prof. Gilles Louppe's Science with AI Lab (SAIL) within the Montefiore Institute (Department of Electrical Engineering and Computer Science). The research group gathers expertise in deep generative models, approximate Bayesian inference, and applications thereof in the physical sciences (incl. particle physics, astronomy and weather science). The group maintains strong collaborations with research institutions worldwide and has access to good computing resources, including a local GPU cluster and national supercomputing facilities. The city of Liège is located in the heart of Europe, within easy reach of Brussels, Paris or Amsterdam. It offers an affordable and high quality of life, with a rich cultural scene and easy access to nature. The university will provide assistance with visa procedures and can help identify housing options.

IV. How to apply?

Send your application by email to Prof. Gilles Louppe (g.louppe@uliege.be), including:

- a 1-page CV highlighting research achievements and technical expertise;
- a list of 2-5 relevant publications with a brief description of your own contribution (optional for candidates to the PhD position);
- a short research statement (1 page max.) outlining your vision and potential contributions to the project;
- links to code repositories of published work;
- names of academic references.

Positions will remain open until they are filled. Applications will be reviewed on a monthly basis.

—

Version of June 4, 2025.