

Training Opportunity for Polish National Trainees

Reference	Title	Duty Station
PL-2023-SCI-SAS	Evaluating Generative Machine Learning for ESA Datalabs	ESAC, Madrid

Overview of the mission:

The European Space Agency maintains a world-leading Science Programme with missions in heliophysics, planetary science, astrophysics and fundamental physics. Its mission is to 'Empower Europe to lead space science'.

The Department for Science and Operations (SCI-S), within ESA's Directorate for Science, host the scientists and engineers that oversee the space missions from study to end of operations; it develops the science operation systems for the missions and operates the missions in space; it archives and curates their data during operations and beyond. Our main objective is to maximise the scientific output of the missions for the benefit of humankind.

The traineeship is placed at the Data Science and Archives Division of the Science and Operations Department and aims to support innovative data science activities with a view to increasing the science return of ESA's science missions through enhancing the science data exploitation, as well as by improving the efficiency of mission science operations. The division is currently developing a platform to support data analytics initiatives for ESA's space science missions. The platform, ESA Datalabs, is focused on improving collaboration and enabling easy access to data for scientific exploitation. However, there are many opportunities to extend its functionality.

Overview of the field of activity proposed:

The purpose of our platform is to facilitate data-driven projects in the space sciences; the integration of machine-learning concepts is of increasing importance in this field given advances in recent years.

In the scope of this traineeship we are looking to investigate the utility generative AI for various use-cases including improved efficiency of systems supporting ESA missions as well as application to challenges in the space sciences. This comprises use-cases in both natural language processing as well as computer vision. As a Polish National Trainee, you will:

- Review the state of the art in research related to generative AI, in particular related to available open source models in the field of natural language processing and computer vision
- Develop prototypes based on (simulated) datasets to demonstrate the feasibility of implementing productive solutions
- Define evaluation frameworks and tools allowing to test prototypes for applicability to ESA mission context
- Consider utilization of on-premise and cloud-based infrastructure (e.g.

- CloudFerro) to mitigate cost-impact of training larger ML-models
- Deploy developed functionality within ESA Datalabs platform
- Present results at internal events as well as machine-learning conferences such as ML in PL

Required education and skills:

- You should have just completed or be in the final year of your Master' s degree in computer science, data science, astrophysics or a related degree
- Good interpersonal and communication skills
- Ability to work in a multi-cultural environment, both independently and as part of a team
- Fluency in English and/or French, the working languages of the Agency