AKPIK 3: Invited Talks

Time: Tuesday 11:00–12:30 Location: KS H C

Invited Talk AKPIK 3.1 Tue 11:00 KS H C

Large Language Models for Research Data Access — •Jutta Schnabel for the KM3NET-ERLANGEN-Collaboration — ECAP, FAU Erlangen-Nürnberg

In a strive to provide ever more interoperable, complex scientific data to a wide range of users across scientific domains, new approaches to lower the boundaries for the access to research data are called for. To address the collection, reduction and analysis of data, as well as its storage, sharing and finding, Large Language Models (LLMs) offer an intriguing opportunity. They can facilitate the transition in research data management from Big Data to Smart Data by increasing the effectiveness of research workflows and the efficiency of data discovery, curation, and reuse especially in the physics domain. This contribution will introduce the uptake of LLMs for data management as pursued in the Physics-LLM project in ErUMData, starting from the application of LLMs in the KM3NeT collaboration.

Invited Talk AKPIK 3.2 Tue 11:30 KS H C

Deep Learning-Based Imaging of MeerKAT Observations — •Kevin Schmitz — TU Dortmund University, Dortumd, Germany — Lamarr Institute for ML & AI, Dortmund, Germany

Modern radio interferometers such as MeerKAT produce vast amounts of data that enable high-resolution imaging of astrophysical sources but also pose significant challenges for image reconstruction. Classical deconvolution algorithms like CLEAN struggle with noise artifacts and limited scalability. The deep-learning framework radionets addresses these issues by combining simulated observations with neural-network-based image reconstruction.
Using the Radio Interferometer Measurement Equation (RIME), realistic simulations model signal propagation and systematic effects in the measured data. These synthetic datasets allow convolutional neural networks to be trained under controlled conditions, enabling quantitative studies of image quality, uncertainty estimates, and computational efficiency.
The framework reconstructs sparse data directly in Fourier space before transforming them into clean sky images, achieving higher positional accuracy and flux recovery than standard tools, especially for wide-field observations containing both compact sources and diffuse emission. I will present its application to real MeerKAT observations, compare the results with classical methods, and illustrate how physics-guided machine learning improves data quality and interpretation in radio astronomy.

Invited Talk AKPIK 3.3 Tue 12:00 KS H C

Towards FAIR fundamental physics: PUNCH4NFDI approach — •Victoria Tokareva[1], Ivan Knežević[2], Harry Enke[3], and Andreas Haungs[1] — [1]Karlsruhe Institute of Technology, Karlsruhe, Germany — [2]GSI Helmholtzzentrum für Schwerionenforschung GmbH, Darmstadt, Germany — [3]Leibniz-Institute for Astrophysics, Potsdam, Germany

Research data management is becoming increasingly important in data-intensive fundamental physics, including astroparticle, particle, nuclear and astrophysics. Growing data volumes and complexity, increasing interdisciplinarity, high costs of data acquisition and processing, and rising expectations for open science and AI-ready digital research outputs are driving these communities toward FAIR (Findable, Accessible, Interoperable, Reusable) data management practices. Key efforts in this direction include reuse and practical harmonisation of semantic artefacts such as metadata schemas, thesauri and controlled vocabularies; making explicit expert knowledge embedded in collaboration-internal data transformations, specialised software tools, and internal data formats (e.g., FITS, ROOT, HDF5); and defining minimal requirements for reusable digital research objects. This talk presents the multi-layer metadata model and the concept and the use-case-based implementations of Digital Research Product, developed within the PUNCH4NFDI (Particles, Universe, NuClei and Hadrons for the German National Research Data Infrastructure, punch4nfdi.de) consortium in the context of a modern FAIR Data landscape and as core elements of the PUNCH4NFDI Science Data Platform.