

German-Russian Astroparticle Data Life Cycle Initiative to Foster Big Data Infrastructure for Multi-Messenger Astronomy

Victoria Tokareva for the GRADLCI collaboration

What. **GRADLCI** is a German-Russian Astroparticle Data Life Cycle Initiative, based on the concept of data life cycle—a pipeline of data curation used for every data unit through the stages: retrieval, preprocessing, analysis, publishing and archival.

Why. Combining data from different experiments is crucial to obtain new knowledge in astroparticle physics; in the framework of GRADLCI we are developing a distributed data center for that purpose.

Results. A prototype of GRADLCI data center has been developed and deployed, providing access to data from KASCADE, KASCADE-Grande, LOPES, Tunka-Rex, Tunka-133, Maked-Ani experiments; new KCDC releases have been delivered, that implement JupyterHub analysis environment, REST API, and extend datashops, simulations, and manuals.

Applications. The data provided by the new infrastructure have been used to develop machine learning techniques for analysis of extensive air showers, and to design labs, masterclasses and exercises on data analysis for various outreach and educational activities.