

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

Friday 16 July 2021 13:18 (12 minutes)

Challenges faced by researchers in multi-messenger astroparticle physics include: computing-intensive search and preprocessing related to the diversity of content and formats of the data from different observatories as well as to data fragmentation over separate storage locations; inconsistencies in user interfaces for data retrieval; lack of the united infrastructure solutions suitable for both data gathering and online analysis, e.g. analyses employing deep neural networks. In order to address solution of these issues, the German-Russian Astroparticle Data Life Cycle Initiative (GRADLCI) project was created. In addition we support activities for communicating our research field to the public.

The approaches proposed by the project are based on the concept of data life cycle (DLC), which assumes a particular pipeline of data curation used for every unit of the data from the moment of its retrieval or creation through the stages of data preprocessing, analysis, publishing and archival. The movement towards unified data curation schemes is essential to increase the benefits gained in the analysis of geographically distributed or content-diverse data.

Within the project, the datasets and user capabilities available for the KASCADE Cosmic-ray Data Center (KCDC) were expanded. As well, an infrastructure for effective astroparticle data storage system and data curation as well as online analysis was developed. Using this infrastructure, first results on deep-learning based analysis were obtained.

The contribution will provide an overview of the accomplished activities and outline possible trends of future developments.

Keywords

data curation, big data, deep learning, multi-messenger astroparticle physics, KCDC

Collaboration

other (fill field below)

other Collaboration

GRADLCI

Subcategory

Experimental Methods & Instrumentation

Primary author: TOKAREVA, Victoria (KIT)

Co-author: FOR THE GRADLCI COLLABORATION

Presenter: TOKAREVA, Victoria (KIT)

Session Classification: Discussion

Track Classification: Scientific Field: MM | Multi-Messenger