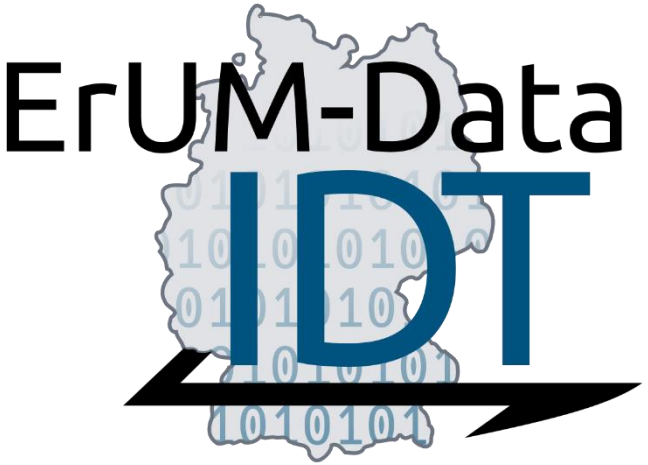
Workshop:

Big Data Analytics: Organisation and Preparation of BMBF Proposal

Andreas Haungs
Markus Roth

11 September 2020



Aktuelle Arbeiten am KIT (Astroteilchenphysik) im Bereich Deep Learning:

- **Simulation of Extensive Air Showers with Deep Neural Networks**
  - Auger, CORSIKA (ErUM-Data Pilot Project)
- **Neural network performances for station-level muon studies**
  - Auger
- **Cosmic Ray composition studies with IceCube Observatory:**
  - IceCube
- **Signal cleaning of radio measurements of extensive air-showers**
  - Auger, IceCube
- **Primary Gamma Ray Search at KASCADE-Grande and Tunka-133**
  - Multi-Messenger

Im Allgemeinen:

- **Simulation und Rekonstruktion von ausgedehnten Luftschauern zur Bestimmung von primärer Teilchenenergie und Teilchenart**

- **Background classification for KATRIN**
  - KATRIN, DARWIN

> Big data analytics: (Gregor)

- > The goal of this project is to develop the necessary tools and environments necessary to master the deluge of data, focusing on activities benefiting more than a single experiment. We foresee the key areas:

1. Tools for analysis and inference on large datasets (distributed training, common inference engine, integration of new hardware architectures into experimental workflows, machine learning as a service,...)
2. Resource efficient generation and simulation (accelerating Monte Carlo simulation, generative ML models, reduction of the simulation gap,...)
3. Improved reconstruction and pattern recognition (tracking, ACATS, realistic and time-dependent conditions,…)
4. Real-time decision making (data processing on hardware, machine learning on FPGAs, triggering, fast event reconstruction,...)

Mögliche Beiträge für neues Projekt

- **Simulation von Luftschauern / Parallelisierung von CORSIKA8**
  - **Beschleunigen von MC (Topic 2)**
- **Analyse von Auger/IceCube/Multi-Messenger Daten**
  - **Pattern recognition (Topic 3, Topic 1)**
- **Background-Unterdrückung (Signalerkennung) und Triggererzeugung bei Radiodaten**
  - **Triggererkennung, auch FPGA (Topic 4)**
- **Background-Klassifikation KATRIN**
  - **(Topic 1, Topic 3)**
- **Analyse von Monitoring Daten bei Virgo**
  - **Realtime Analyse (Topic 1, Topic 4)**
- **….**
  - **…**

➔ **Details zu definieren unter Einbeziehung des Gesamtkonzeptes für diese Antragsidee**