AKPIK 7: AI Topical Day – Research Data Management and Medical Applications

Time: Thursday 14:00–15:30 Location: HSZ/0101

AKPIK 7.1 Thu 14:00 HSZ/0101

Federated Heterogeneous Compute and Storage Infrastructure for the PUNCH4NFDI Consortium — •Alexander Drabent[1], Jörn Künsemöller[2], Matthias Hoeft[1], Christoph Wissing[3], Manuel Giffels[4], Dominik Schwarz[2], Kilian Schwarz[3], and Andreas Henkel[5] — [1]Thüringer Landessternwarte Tautenburg — [2]Universität Bielefeld — [3]DESY, Hamburg — [4]Karlsruher Institut für Technologie — [5]Johannes Gutenberg-Universität Mainz

PUNCH4NFDI is the NFDI consortium of particle, astro-, astroparticle, hadron and nuclear physics.

Compute4PUNCH and Storage4PUNCH concepts are developed to meet the diverse needs of these communities to provide seamless and federated access to compute and storage systems. Those are being federated in a common infrastructure and transparently integrated with an overlay batch system. Both concepts comprise state-of-the-art technologies for resource access and to ensure scalable provisioning of community specific software. Furthermore, existing technologies for caching as well as metadata handling are being evaluated with the aim for a deeper integration. The combined Compute4PUNCH and Storage4PUNCH environment will allow a large variety of researchers to carry out resource-demanding analysis tasks.

In this contribution we will present the Compute4PUNCH and Storage4PUNCH concepts, the current status of the developments as well as first experiences with scientific applications, such as analysing radio-interferometric data, being executed on the available prototypes.

AKPIK 7.2 Thu 14:15 HSZ/0101

VISPA - Cloud Services for Modern Data Analysis — •Niclas Eich, Louis Christoph, Martin Erdmann, Peter Fackeldey, Benjamin Fischer, Leonard Lux, Dennis Noll, Mathilde Pöppelmann, and Malcom Steen — RWTH Aachen University

VISPA (VISual Physics Analysis) realizes a scientific cloud enabling modern scientific data analysis in a web browser. Our local VISPA instance is backed by a small institute cluster and is dedicated to fundamental research and university education. By hardware upgrades (656 CPU threads, 29 workstation GPUs), we have tailored the cloud services to accomplish both, rapid turn-around when developing O(TB) HEP analyses and deep-learning hands-on with O(100) participants through the web browser. With its latest software developments, VISPA now supports the interactive use of Jupyter notebooks on local as well as on batch resources. Additionally, users can optionally execute their analyses on any SSH reachable large-scale resource they desire. New tools such as an improved user management and a monitoring of the batch resources ensure seamless administration. We will present this major advance of the VISPA project and show how a wide range of scientific data analyses can be realized in the web browser.

AKPIK 7.3 Thu 14:30 HSZ/0101

Towards coherent metadata schema for the PUNCH4NFDI open science platform — •Victoria Tokareva for the PUNCH4NFDI Consortium-Collaboration — Karlsruhe Institute of Technology, Institute for Astroparticle Physics, Karlsruhe, Germany

PUNCH4NFDI is an NFDI consortium of particle, astro-, astroparticle, hadron and nuclear physics, which is addressing common challenges of data-intense physics at large research facilities: data volumes, data complexity, data rates, and data irreversibility, as well as the development and promotion of open science vision and required tools to achieve this. One of the core features in the development of a PUNCH4NFDI software infrastructure is a cloud-based platform and an open data portal, aimed at providing access to a wide range of digital research materials within the PUNCH4NFDI community and ensuring that the FAIR (findability, accessibility, interoperability, reusability) principles are applied for the community's data collections. This requires to navigate the landscape of different established metadata schemas and find common ground to access the data and run programs and workflows using data from different data collections. In order to achieve this goal, we have investigated core concepts and definitions in the field and analyzed user stories and use cases of several data platforms within the PUNCH4NFDI community. From these, essential requirements for the used metadata have been defined. The results will be presented in this contribution. This work is supported by the DFG fund "NFDI 39/1" for the PUNCH4NFDI consortium.

AKPIK 7.4 Thu 14:45 HSZ/0101

Datenanalytische Hilfestellungen für ein festgelegtes Modell zur Personenerkennung — •Jan Michael Bürger und Hans Dominik Werner — HowRyou GmbH, 24976 Handewitt

Für den Bereich der (Alten-)Pflege werten wir Videodaten mit einer Personenerkennung aus, um eine Videokommunikation genau dann zu ermöglichen, sobald sich die Person alleine im Raum aufhält. Insbesondere aus Datenschutzgründen sollte dabei eine Kommunikation nicht möglich sein, sobald sich mehr als eine Person im Raum aufhält. Gleichzeitig sollte die Person erkannt werden, selbst wenn sie zugedeckt im Bett liegt.

Für diese Aufgabe greifen wir auf bereits verfügbare Modelle zur Personenkennung zurück. Um unsere Anforderungen bestmöglich zu erfüllen, wäre es zunächst naheliegend, dass Modell anzupassen bzw. dieses mit geeigneten Trainingsdaten nachzutrainieren. Um vor allem mit einem kleineren Datensatz auszukommen, haben wir den Fokus auf einen anderen Ansatz gelegt: Wir haben eine systematische Datenanalyse der Personenerkennung auf Testdaten durchgeführt.

Auf Grundlage der Ergebnisse dieser Datenanalyse haben wir (einfache) Techniken implementiert, die die Videobilder im Vorfeld graphisch manipulieren und die Ergebnisse auf geeignete Weise verrechnen. Dazu zählt u.a. das Präsentieren des gleichen Bildes in verschiedenen Helligkeiten und ein virtuelles Drehen des Bildes. Diese Hilfestellungen führten zu zum Teil signifikanten Verbesserungen.

In diesem Vortrag werden die Erfahrungen und Ergebnisse zu von uns untersuchten und verwendeten Hilfestellungen dargestellt.

AKPIK 7.5 Thu 15:00 HSZ/0101

Interpretable Machine Learning and evidence-based decision support in clinical Digital Twins — •Carlos Andres Brandl, Anna Nitschke, and Matthias Weidemüller — Im Neuenheimer Feld 226, 69120 Heidelberg, Germany

Personalized medicine is based on including a vast variety of patient-specific data. The Digital Twin technology provides the opportunity for improved personalized patient care by monitoring the patient journey and predicting the best preventive and therapeutic decision options available. We developed a concept which fuses evidence-based methods with machine learning approaches into a single decision-support tool. Our method is independent on the parameter spaces and evidence-based tools being used, provides possibilities to include updated knowledge and is able to offer intuitively interpretable decision options to the clinician. The presentation introduces our architecture of the digital twin and provides details on the fusion approach.

AKPIK 7.6 Thu 15:15 HSZ/0101

Radiomics for two-dimensional prompt gamma-ray based proton treatment verification — •Sonja M Schellhammer[1,2], Theresa Lenk[1,3], Steffen Löck[1,3,4], and Toni Kögler[1,2] — [1]OncoRay – National Center for Radiation Research in Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Helmholtz-Zentrum Dresden - Rossendorf, Dresden, Germany — [2]Helmholtz-Zentrum Dresden - Rossendorf, Institute of Radiooncology – OncoRay, Dresden, Germany — [3]German Cancer Consortium (DKTK), Partner Site Dresden, and German Cancer Research Center (DKFZ), Heidelberg, Germany — [4]Department of Radiotherapy and Radiation Oncology, Faculty of Medicine and University Hospital Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

There is a growing need for on-line verification systems to further increase the safety and efficacy of cancer treatment with proton therapy. For this purpose, we propose a radiomics-based analysis of two-dimensional time-energy distributions of secondary prompt gamma-rays and apply it to realistic data measured in a proton therapy facility. In comparison to previously used methods based only on temporal gamma-ray distributions, we show that the accuracy of the range verification is improved by 38 % (1.5 mm). These results demonstrate that radiomics and machine learning are valuable tools to enhance proton treatment verification for cancer therapy.