Proposal for input to ESPP update in the area of computing, software, ML

The full exploitation of the unique physics potential of the to be collected data sets in the Exabyte regime requires the development and application of state-of-the-art software and machine learning methods and the provisioning of extended compute and storage resources beyond the increase provided by technological evolution. At the same time scientific computing has to become CO2-neutral by 2050 in order to mitigate the impact on climate change and to meet the agreed on climate goals within the European Union. Hence, a timely development of novel computing concepts and innovative algorithms for data handling, event generation, detector simulation, reconstruction and analysis taking into account the criteria of the magnitude of the to be collected data and of the required compute power and the criteria of environmental sustainability is imperative.

WLCG, using mainly dedicated hardware resources, has been a great success over the last two decades. Nevertheless, the option of an interdisciplinary usage of large IT infrastructures (e.g. NHR[§] in Germany, EOSC at European level) shall be investigated and be carefully evaluated. Necessary software frameworks for orchestration and accounting shall be developed. The full costs for provisioning and operation of the required hardware resources shall be included in the financial planning of future experiments.

Modern AI concepts are essential for event generation, detector simulation, event reconstruction and data analysis. A flexible common experiment overarching software ecosystem for all these tasks, such as the turnkey software-stack KEY4HEP, is vital for the development and optimisation of the next generation of detectors at future colliders. It is mandatory that the seamless and easy integration of modern AI algorithms is provided from the start. The development and maintenance of new concepts and software libraries for the above tasks shall continuously be supported.

Interdisciplinary cooperation on national (e.g. NFDI, DIG-UM[§] in Germany) and international level (e.g. WLCG, JENA, EOSC) shall be continued and further be enhanced.

The data and research results of the particle physics community create a unique legacy. Hence, it is mandatory to develop technical solutions for sustainable Open Science, Data Preservation and scientific reuse strategies in accordance with FAIR principles.

To be able to profit from future technological developments in computing, software and data analysis, which may be fundamental and disruptive, it is essential to keep the relevant expertise in the community. This can only be achieved by a recognition of work in this area, in particular conducted by experimental early career researchers, at the same level as for R&D in detector and accelerator technologies and for data analysis.

[§]NHR = National High Performance Computing Alliance
[&]NFDI = National Research Data Infrastructure, DIG-UM = Digital Transformation in the Research of Universe and Matter