

The SMARTHEP European Training Network

James Andrew Gooding^{1,}, Author 2^{2,**}, Author 3^{2,***}, Author 4^{2,****}, Author 5^{2,†}, Author 6^{2,‡}, and Author 7^{2,§}*

¹Fakultät Physik, Technische Universität Dortmund, Dortmund, Germany

²Institute 2

³Institute 3

⁴Institute 4

⁵Institute 5

⁶Institute 6

⁷Institute 7

Abstract. Synergies between **MA**chine learning, **Re**al-**T**ime analysis and **Hy**brid architectures for efficient **E**vent **P**rocessing and decision-making (SMARTHEP) is a European Training Network, training a new generation of Early Stage Researchers (ESRs) to advance real-time decision-making, driving data-collection and analysis towards synonymity.

SMARTHEP brings together scientists from major LHC collaborations at the frontiers of real-time analysis (RTA) and key specialists from computer science and industry. By solving concrete problems as a community, SMARTHEP will further the adoption of RTA techniques, enabling future HEP discoveries and generating impact in industry.

ESRs will contribute to European growth, leveraging their hands-on experience in machine learning and accelerators towards commercial deliverables in fields that can profit most from RTA, e.g. transport, manufacturing, and finance.

This contribution presents the training and outreach plan for the network, and is intended as an opportunity for further collaboration and feedback from the CHEP community.

1 Introduction

The Synergies between **MA**chine learning, **Re**al-**T**ime analysis and **Hy**brid architectures for efficient **E**vent **P**rocessing and decision making (SMARTHEP) European Training Network is an EU Horizon-funded training network, with a focus on the development of expertise in real-time analysis (RTA) techniques through applications to high energy physics (HEP) research and industry. The network centres around the training of 12 Early Stage Researchers (ESRs) between September 2021 and September 2025.

*e-mail: Jamie.Gooding@cern.ch

**e-mail: a.2@cern.ch

***e-mail: a.3@cern.ch

****e-mail: a.4@cern.ch

†e-mail: a.5@cern.ch

‡e-mail: a.6@cern.ch

§e-mail: a.7@cern.ch

2 SMARTHEP as a European Training Network

As a European Training Network (ETN), the primary aim of the network is in training ESRs, whilst deepening synergies between HEP and industry. The network takes a novel approach to building such synergies, structuring each ESR position (a 3 year period of doctoral study) around academic and industrial secondments. To achieve this, the network is formed of a series of partnerships between universities, research institutes and organisations in industry, as listed in Table 1.

Category	Partners
Universities	Lund University, Sorbonne University (LPNHE & LIP), Technische Universität Dortmund, University of Bologna, University of Geneva, University of Heidelberg, University of Helsinki, University of Manchester, Universidade Santiago de Compostela (IGFAE), Vrije Universiteit Amsterdam
Research institutes	CERN, CNRS, NIKHEF
Industry partners	IBM France, Lightbox, Point 6, University of Manchester Institute for Data Science and AI, Verizon Connect, Ximantis

Table 1. The 18 organisations which form the SMARTHEP network.

The network is formally structured as 7 Work Packages (WPs), laid out in Figure 1. WP1 “Management”, covers the management of the network by the Project Manager/Project Coordinator/Executive Board. WP2 “Training”, sets out training and development of participants, by both network partners and external training providers. WP3 “Machine learning & advanced data analysis”, develops machine learning (ML) techniques for use in real-time environments. WP4 “Hybrid architectures” focuses upon the deployment of non-CPU computing architectures in acceleration of data processing. WP5 “Decision-making in research and industry”, applies WP3 and WP4 to develop RTA-based decision-making technologies. WP6 “Monitoring and discoveries”, also applies WP3 and WP4 to RTA approaches in data analysis. WP7 “Dissemination and communication of results”, publishes and propagates the results produced in WP6 and WP7.

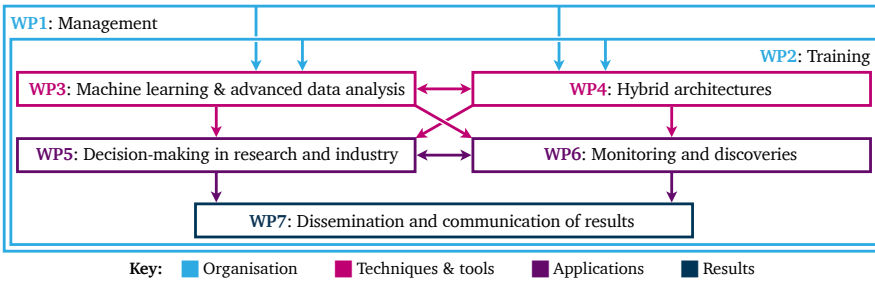


Figure 1. The structure of work packages within the SMARTHEP network. WP1 and WP2 define the organisation of the network; WP3 and WP4 introduce the techniques and tools of real-time analysis to the network; WP5 and WP6 use said techniques and tools to produce results for HEP and industry; WP7 makes these results available and promotes their wider use and adoption.

The network has a particular focus on physics at the Large Hadron Collider (LHC). Each ESR is thus affiliated to one of the four major experiments based at the LHC: ALICE, ATLAS, CMS and LHCb. The network duration coincides with Run 3 of the LHC (2022-2025), with many of the ESR projects framed in this context.

A unique feature of the network is the extensive cooperation between HEP and industry across all ESR positions. RTA approaches have seen significant adoption in industry in recent years, with many organisations turning to RTA as a means to handle “big data”.

3 Real-time analysis

HEP and industry share a common challenge—the rapid processing of large quantities of data. [1] Recent advances in computing, in particular in the areas of machine learning and hybrid architectures, have enabled the possibility of processing data in real-time, i.e. as data is collected (also commonly referred to as “online” processing). [2] By processing data online, resources (computing power, storage space, energy, etc.) can be saved and further insights can be obtained from the data recorded, as shown in Figure 2.

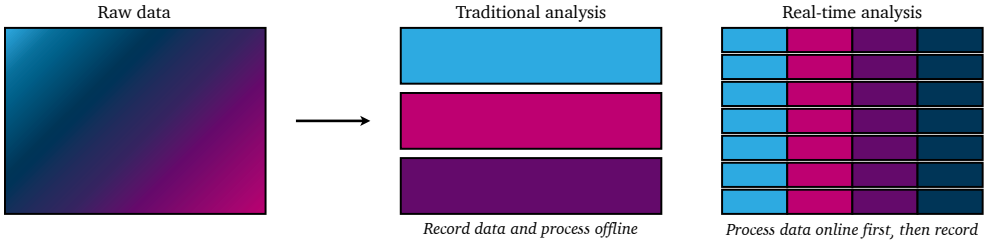


Figure 2. Traditional and RTA approaches to data processing. Traditional approaches rely on recording all data and processing this offline; in RTA, data is processed as it is produced, recording only the relevant portions, enabling greater volumes of processed data to be stored.

RTA techniques have seen widespread adoption across HEP, in particular amongst trigger and data acquisition (TDAQ) systems. Since it is not possible to record a full detector readout and carry out event reconstruction at the LHC collision rate of 40 MHz, triggers must be devised to select only those events relevant to the physics goals of an experiment. Such triggers conventionally consist of a hardware system making coarse decisions from partial detector readout, and a staged software trigger, applying gradually more finely-grained selections with increasingly more detailed reconstructions of events.

In industry, the limitations of the scale of “big data” looms over many applications of computing. However,

3.1 Machine learning

ML, a catch-all term for a family of techniques and technologies wherein algorithms are trained to analyse data, enables rapid decision-making and pattern recognition across a broad range of use cases. [3] In HEP, the adoption of ML began with classifiers for offline physics analysis, since widening to a variety of classification and pattern recognition/anomaly detection techniques for use online and offline. [4]

ML classifiers are algorithms classed with classification, with Boosted Decision Trees (BDTs) and Neural Networks (NNs) amongst those most commonly employed in HEP. The use of BDTs in signal selection for offline physics analysis has become standard practice, e.g. suppression of combinatorial background in the LHCb experiment measurement of the $B_s^0 - \bar{B}_s^0$ oscillation frequency. [5] Classifiers can also aid tasks such as event reconstruction, e.g. the CMS boosted event shape tagger, a NN trained to discriminate between possible t ,

W^\pm , Z^0 and H candidates within an event. [6] In industry, ML classifiers are used frequently in content deliver

ML techniques can also be applied to pattern recognition and anomaly detection—tasks which cannot be realistically carried out on the scales of data presently being analysed. In industry, this is applied intuitively to fraud detection, wherein even subtle changes in transaction data can indicate fraudulent activity. In HEP, such anomaly detection can be applied to searches, for example the ATLAS search for resonant decays to a Higgs boson. [7]

3.2 Hybrid architectures

Typical computational resources consist of Central Processing Units (CPUs), processors which are designed for general purpose computation (i.e. varied tasks with significant variety in computing/memory resources), with large on-board memory and often multiple processing cores. Alternative computing architectures, such as Graphical Processing Units (GPUs) and Field-programmable gate arrays (FPGAs) shown in Figure 3, can be applied in conjunction or in place of CPU architectures to accelerate processing where tasks are not well-suited to CPU architectures alone.

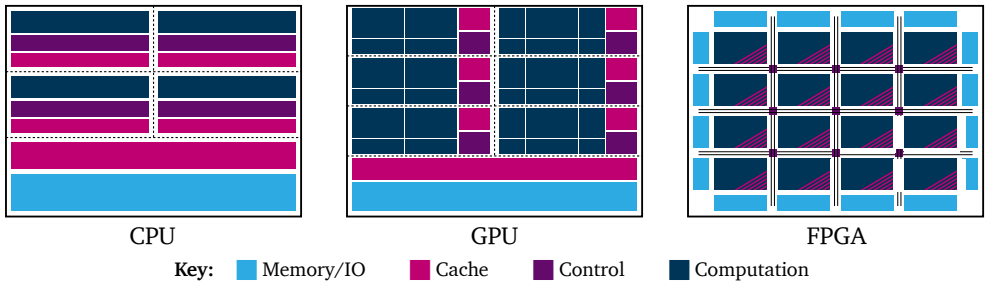


Figure 3. Comparison of CPU, GPU and FPGA architectures. A CPU is typically formed of several cores (each containing computational, control and cache resources) and centralized cache, memory and input/output (IO) interface. GPUs contain similar centralized resources, though consist of many multiprocessors, each containing a greater proportion of computational resources than a CPU core. GPU multiprocessors are also themselves partitioned to perform tasks in parallel. FPGAs are structured rather differently with memory and IO connected to many interlinked control blocks formed of simpler logic gate arrangements, often accompanied by a small cache.

GPUs typically contain many more computing units than CPUs, providing better performance in highly parallelisable tasks. This makes GPUs well-suited to accelerating event processing, which consists of many similar but computationally intensive tasks such as cluster finding and track fitting. [8] For less computationally intensive tasks, FPGAs (integrated circuits which can be programmed for specific tasks) can be employed. [9] Examples of such tasks include low-level trigger decisions, which typically consist of simple selections but must be made rapidly. [10]

4 Early stage researchers

12 ESRs form the core of the network, with the training and partnerships providing a scaffold for the completion of their respective outcomes. Each ESR is enrolled as a doctoral student at a partner university for 3 years, during which they complete secondments in HEP and industry, as illustrated in Figure 4. The university of enrolment is generally also the primary

institution of the ESR (industry-centred ESR positions are enrolled at a university near to the respective industry partner). A secondment in HEP is undertaken either at another partner university or a partner research institute. Each ESR (with the exception of those in industry-centred positions) also undertakes a secondment in industry, working on an RTA project relevant to their research at an industry partner.

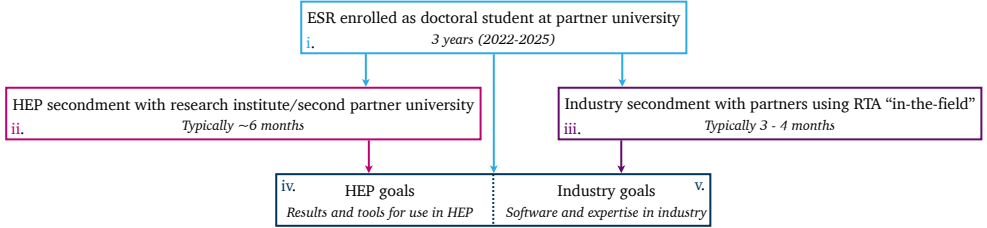


Figure 4. Structure of a SMARTHEP ESR position. Each ESR is enrolled (i.), during which they will undertake secondments with network partners in HEP (ii.) and industry (iii.). Through the combination of primary and secondment work, each ESR will complete goals in HEP (iv.) and industry (v.), discussed in further detail in Section 5.1. Precise durations of the secondments vary between ESR positions.

To illustrate further the structure of a SMARTHEP ESR position, a few examples are given:

- Laura Boggia, one of two industry-centred ESRs, is researching fraud detection at IBM (enrolled at Sorbonne University) using real-time rule induction. A secondment with the CNRS research institute is planned, in which similar techniques will be applied to the classification of HEP observations.
- Carlos Cocha is based at the University of Heidelberg, working on real-time searches for Dark Photons at LHCb, furthered by collaboration with the University of Milano-Bicocca. An industry secondment with Verizon Connect is planned, applying ML expertise to the real-time processing of vehicle data.
- Joachim Carlo Kristian Hansen is developing the real-time calibration of the ALICE Time Projection Chamber, working at Lund University. This commissioning work will be supported by a secondment at CERN. An additional secondment at Ximantis will promote a synergy between industry and HEP applications of ML.

5 Network outcomes

The network is defined around a set of concrete outcomes, guiding the progression of the network and its participants. These intended outcomes are summarised briefly below.

5.1 Goals of HEP and industry

The core outcomes of the network are a set of goals, completed on behalf of the network partners by the network participants. These goals include experiment commissioning, HEP measurements and industrial results. These goals are typically defined with respect to a specific ESR position, e.g. the goal of “calibration of ALICE TPC for heavy-ion physics” associated to Joachim Hansen.

5.2 RTA whitepapers

Participants will write a series of three whitepapers reviewing the current RTA state-of-the-art, expected to be submitted in late 2023. These whitepapers will place a particular emphasis on the ongoing contributions of the network to their respective topics.

The first whitepaper reviews ML applications to HEP in RTA contexts and their corresponding best practices. Such applications range from data-taking, e.g. particle identification at ALICE, to offline analysis, e.g. anomaly detection for dijet resonance searches at ATLAS. [11, 12] The use of hybrid architectures by LHC experiments is reviewed in the second whitepaper. Tasks such as selection and reconstruction were significantly accelerated during Run 2 of the LHC (2015-2018). Run 3 deployments of hybrid architectures capitalise on this progress, e.g. in the use of FPGAs in the ALICE Central Trigger System and the use of GPUs in the ATLAS High Level Trigger farm. [13, 14] In the third whitepaper, TDAQ systems of LHC experiments are reviewed, with a focus upon best practices for both TDAQ hardware and software. As such, this whitepaper reviews topics ranging from upgrades to the ATLAS TDAQ system, to the Allen framework of the LHCb experiment enabling software trigger operation at a 30 MHz readout rate. [15, 16]

5.3 Training of ESRs

As a key network objective, participants are encouraged to partake in a broad range of training activities, with network resources dedicated to facilitating this. Such activities include the attendance of external specialist industrial training sessions and academic schools, in addition to internally organised workshops and schools (e.g. in Section 6.1).

5.4 Software and digital assets

The work of the ESRs will generate digital assets (e.g. software packages, data processing tools), with many being applicable beyond narrow academic/industrial applications. The network is therefore committed to making any such digital assets Findable, Accessible, Interoperable, and Reusable (FAIR). [17] To implement these commitments, a project on GitHub has been created to host such assets, <https://github.com/SMARTHEP>. Other assets and resources will be made available on the network website, <https://smarthep.org>.

6 Network events

Network events provide participants with unique opportunities to meet, exchange ideas and develop. Whilst these events typically cater to the ESR audience, some events have been made available to additional interested early career researchers working/studying at SMARTHEP institutes. Previous and upcoming network events are briefly summarised below.

6.1 Previous events

To commence ESR participation in the network, a kick-off meeting was held at the University of Manchester in November 2022 to formally introduce the ESRs to the network and discuss network objectives and organisation. Amongst the activities of the kick-off meeting were a visit to the Jodrell Bank Centre for Astrophysics and a review paper-writing course by Scriptoria to train ESRs ahead of writing the whitepapers.

In January of 2023, the First SMARTHEP School on Collider Physics and Machine Learning was hosted by the University of Geneva. The school provided a varied programme, including lectures on experimental physics at collider experiments by Anna Sfyrla and on theoretical physics & Monte Carlo event generators by Torbjörn Sjöstrand. Additionally, hands-on lessons in machine learning were led by Maurizio Pierini, with seminars also given on multimessenger astronomy and the CERN experimental programme by Teresa Montaruli and Jamie Boyd respectively.

6.2 Upcoming events

Upcoming network events aim to guide the governance of the network, develop ESR expertise and deepen collaborations between participants. The network assembly serves as an annual forum in which to make significant decisions on network policy and governance. Accelerator and ML bootcamps, likely to take place in summer 2024, will formalise the practical experience of ESRs. Collaboration with industry will culminate in an industry applications school late in the network, by which time it is expected that all ESR industry secondments will have been completed.

7 Conclusion

Through the activities of the network outlined in this contribution, SMARTHEP aims to provide valuable contributions to HEP, industry, and the wider community. By the close of the network in September 2025, the 12 ESRs will be well-equipped to progress their research and their careers. The collaboration of academic and industrial partners in the network will deepen synergies across the field in the application of real-time analysis.

Acknowledgements

We acknowledge funding from the European Union Horizon 2020 research and innovation programme, call H2020-MSCA- ITN-2020, under Grant Agreement n. 956086

References

- [1] H. Hu, Y. Wen, T.S. Chua, X. Li, IEEE Access **2**, 652 (2014)
- [2] W. Tian, Y. Zhao, in *Optimized Cloud Resource Management and Scheduling*, edited by W. Tian, Y. Zhao (Morgan Kaufmann, Boston, 2015), pp. 17–49, ISBN 978-0-12-801476-9
- [3] J. Jovel, R. Greiner, Frontiers in Medicine **8** (2021)
- [4] K. Albertsson, P. Altoe, D. Anderson, J. Anderson, M. Andrews, J.P.A. Espinosa, A. Aurisano, L. Basara, A. Bevan, W. Bhimji et al., *Machine learning in high energy physics community white paper* (2019), 1807.02876
- [5] R. Aaij, C.A. Beteta, T. Ackernley, B. Adeva, M. Adinolfi, H. Afsharnia, C.A. Aidala, S. Aiola, Z. Ajaltouni, S. Akar et al., Nature Physics **18**, 1 (2022)
- [6] A. Sirunyan, A. Tumasyan, W. Adam, F. Ambrogio, T. Bergauer, M. Dragicevic, J. Erö, A.E.D. Valle, M. Flechl, R. Frühwirth et al., Journal of Instrumentation **15**, P06005 (2020)
- [7] (2023), 2306.03637
- [8] D. vom Bruch, Journal of Instrumentation **15**, C06010 (2020)

- [9] J. Serrano, *Introduction to FPGA design* (2008), <https://cds.cern.ch/record/1100537>
- [10] J. Duarte, P. Harris, S. Hauck, B. Holzman, S.C. Hsu, S. Jindariani, S. Khan, B. Kreis, B. Lee, M. Liu et al., *Computing and Software for Big Science* **3** (2019)
- [11] Ł.K. Graczykowski, M. Jakubowska, K.R. Deja, M. Kabus, on behalf of the ALICE collaboration, *Journal of Instrumentation* **17**, C07016 (2022)
- [12] G. Aad, B. Abbott, D.C. Abbott, A. Abed Abud, K. Abeling, D.K. Abhayasinghe, S.H. Abidi, O.S. AbouZeid, N.L. Abraham, H. Abramowicz et al. (ATLAS Collaboration), *Phys. Rev. Lett.* **125**, 131801 (2020)
- [13] Kvapil, Jakub, Bhasin, Anju, Bombara, Marek, Evans, David, Jusko, Anton, Kluge, Alexander, Krivda, Marian, Kralik, Ivan, Lietava, Roman, Nayak, Sanket Kumar et al., *EPJ Web Conf.* **251**, 04022 (2021)
- [14] A. Bocci, on behalf of the CMS Collaboration, *Journal of Physics: Conference Series* **2438**, 012016 (2023)
- [15] G. Aad, T. Abajyan, B. Abbott, J. Abdallah, S. Abdel Khalek, O. Abdinov, R. Aben, B. Abi, O. AbouZeid, H. Abramowicz et al. (ATLAS), *Tech. rep.* (2013), final version presented to December 2013 LHCC., <https://cds.cern.ch/record/1602235>
- [16] R. Aaij, J. Albrecht, M. Belous, P. Billoir, T. Boettcher, A. Brea Rodríguez, D. vom Bruch, D.H. Cámpora Pérez, A. Casais Vidal, D.C. Craik et al., *Computing and Software for Big Science* **4**, 7 (2020)
- [17] M.D. Wilkinson, M. Dumontier, I.J. Aalbersberg, G. Appleton, M. Axton, A. Baak, N. Blomberg, J.W. Boiten, L.B. da Silva Santos, P.E. Bourne et al., *Scientific Data* **3**, 160018 (2016)