START PAGE

MARIE SKŁODOWSKA-CURIE ACTIONS

Innovative Training Networks (ITN)
Call: H2020-MSCA-ITN-2019

Synergies between Multivariate
Analysis, Real Time analysis and Hybrid architectures for Event Processing
SMARTHEP

This proposal is to be evaluated as:
ETN

LIST OF PARTICIPATING ORGANIZATIONS

List of Beneficiaries

Consortium Member (beneficiaries)	Legal Entity Short Name	Academic	Non-Academic	Awards Doctoral Degrees	Country	Dept. Division Laboratory	Scientist-in-Charge
1. Lund University	ULUND	✓		✓	SE	Particle Physics	Assistant Prof. Caterina Doglioni
2. CNRS	CNRS	✓			FR	LPNHE	Dr. Vladimir Gligorov
3. TU Dortmund University	TUDO	✓		✓	DE	Faculty of Physics, E5	Dr. Johannes Albrecht
4. Heidelberg University	UHEI	✓		✓	DE	UHEI	Prof. Stefanie Hansmann-Menzemer
5. University of Helsinki	UH	✓		✓	FI	Department of Physics	Assistant Prof. Mikko Voutilainen
6. EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH	CERN	✓			CH	Physics Department	Dr. Monica Pepe Altarelli
7. National Institute for Subatomic Physics, Nikhef	NIKHEF	✓			NL	Scientific Department	Prof. Olga Igonkina
8. Université de Genève	UNIGE	✓		✓	CH	DPNC	Associate Prof. Anna Sfyrla
9. Sorbonne Université	SU	✓		✓	FR	LIP6	Prof. Lionel Lacassagne
10. IBM France	IBM		✓		FR	–	Dr. Christian de Sainte Marie
11. KKT	KKT		✓		IT	–	Dr. Francesco Sambo

Data for non-academic beneficiaries

Name	Location premises	Type of R&D activities	Full-time employees	Employees in R&D	Website	Annual turnover	Enterprise status	SME status
KKT (subsidiary of Verizon)	Firenze, IT	fleet intelligence	35	34	`https://www.verizonconnect.com`	109 B€ (consolidated with Verizon)	YES	NO
IBM	Paris, FRA	algorithm and software development	7500	600	`https://www.ibm.com/fr-fr/`	2,15 B€	YES	NO

FTE = Full-time equivalent

List of Partner Organisations

Consortium Member (Partner)	Legal Entity Short Name (also used in this Document)	Academic	Non-Academic	Awards Doctoral Degrees	Country	Dept. Division Laboratory	Scientist-in-Charge	Role of Partner Organisation
Ximantis	XIMANTIS		✓		SE	–	Dr. Alexandros Sopasakis	Training, secondments on RTA in traffic prediction and decision-making
POINT8	POINT8		✓		DE	–	Dr. Kevin Dungs	Training, secondments on RTA in transport and logistics
Lightbox	LIGHTBOX		✓		IT	–	Dr. Pierluigi Catastini	Training, secondment on RTA in IoT in industrial processes and financial decisions (based in Geneva)
University of Pisa and INFN	INFN	✓		✓	IT	Department of Physics	Prof. Chiara Roda	Training, secondments in hybrid computing architectures for ATLAS
University of Santiago de Compostela	USC	✓		✓	ES	Department of Physics	Ass. Prof. Diego M. Santos	Training, secondments in GPU for RTA and physics for LHCb
University of Oregon	UO	✓		✓	US	Physics Department	Prof. David Strom	Training, secondment at CERN in hybrid computing architectures and physics for ATLAS
The Ohio State University	OSU	✓		✓	US	CCAPP and Physics Department	Ass. Prof. Antonio Boveia	Training, secondment at CERN in FPGA, hybrid computing architectures and physics analysis for ATLAS
University of Liege	ULIEGE	✓		✓	BE	Department of Physics	Ass. Prof. Gilles Louppe	Training, secondment in machine learning for anomaly detection
VU University Amsterdam	VU	✓		✓	NL	Department of Physics	Prof. Gerhard Raven	Training, award of PhD degree
Radboud University Nijmegen	RU	✓		✓	NL	Department of Physics	Mario C. van der Toorn	Training, award of PhD degree
University of Bologna	UNIBO	✓		✓	IT	Department of Computer Science	Dr. Samuele Salti	Training, award of PhD degree

We declare no inter-relationship between different participating institutions or individuals.

B1. EXCELLENCE

The volume of data available to research and industry is ever-increasing. This acceleration in data collection is not always matched by a comparable acceleration in data storage and data utilization. This poses a **problem for the traditional paradigm of first collecting data, then analyzing it**, in both research and industry. In cases when the space to store the data is limited, then data has to be discarded or not recorded at all. Moreover, most of the data produced nowadays is recorded and stored without ever being analyzed, and loses value if insight is not extracted in a short period of time[1].

For High Energy Physics (HEP) experiments at the Large Hadron Collider (LHC) at CERN, proton-proton collisions occur up to 30 million times per second. A decision on whether a collision "event" is interesting must be made in the order of milliseconds, and novel techniques are needed in order to extract information that would otherwise be discarded. Commercial applications such as mobile apps for fleet control and transport optimization, Internet-of-Things (IoT) sensors and fraud detection face a similar challenge: the amount and complexity of the data grows while the resources and time to take a decision using that information do not scale accordingly.

In order to make the most of the available information in a cost-effective way, and to reduce the time between data collection and insight on results, data taking and data analysis need to be fast and efficient in both HEP and industry. SMARTHEP (**S**ynergies between **M**ultivariate **A**nalysis, **R**eal **T**ime analysis and **H**ybrid architectures for **E**vent **P**rocessing) is an interdisciplinary ETN aiming to train Early Stage Researchers (ESRs) on **Real Time Analysis (RTA)** techniques, where **data collection and analysis are nearly simultaneous**, using software and hardware tools beyond the state-of-the-art. Within SMARTHEP, the ESRs will employ machine learning (ML) and multivariate techniques to speed up data reconstruction, selection and analysis, introducing novel algorithms that improve reproducibility and accountability of real-time decision making. The ESRs will also become experts in novel hybrid software/hardware solutions, where data can be collected and reconstructed on the same platform in order to reduce the time-to-insight.

The SMARTHEP **ESRs will be trained** by a consortium of successful researchers in **High Energy Physics and Computer Science** (including 6 holders of ongoing ERC StG and CoG and recipients of multiple national grants, as well as a mentoring group of senior experienced scientists worldwide) and entrepreneurs from industries ranging from start-up to multinational. These scientists, professionals and entrepreneurs have begun their collaboration within SMARTHEP and will continue working together beyond this action with a commitment to follow the future careers of the scientists they train. The ESRs will **learn transferrable skills** from the exposure to academia and industry in the dissemination and exploitation of their research, as well as to stakeholders and general public while communicating their results. They will be **prepared for careers in industry or science** and **contribute to European growth** from their hands-on work on concrete commercial and research deliverables, such as mobile apps using ML for vehicle fleet monitoring and safety, optimized algorithms for traffic and public transport predictions, software for sensors that will monitor industrial production, accountable and understandable algorithms for fraud detection and decision-making in finance, as well as cutting edge computing technologies and algorithms necessary for the detection and measurement of fundamental particles and for the discovery of new physics phenomena.

B1.1 Quality, innovative aspects and credibility of the research programme

B1.1.1 Introduction, objectives and overview of the research program

Researchers in HEP study the basic constituents of matter and their interactions within the theory called the Standard Model (SM), and search for new physics phenomena explaining the SM's shortcomings. One of the main experimental tools available for this study is the LHC. At the LHC, particle beams are accelerated to high energies and collision *events* occur 30 million times per second. The products of the collisions are analyzed by scientists operating four main particle detectors: ALICE, ATLAS, CMS and LHCb[2]. Recording each event would require 1000 times the available storage, and in many cases over 10 times the available data acquisition bandwidth. Each experiment therefore has a *trigger system* composed of hardware and software. The trigger system analyzes collision events in real time, with a maximum delay of micro- to milliseconds, and selects the ones most interesting for further analysis. While there are many observed phenomena that are still not included in the SM (e.g. dark matter [3]), the leading predictive theories such as Supersymmetry have not yet yielded a discovery. This is prompting HEP researchers to reconsider their definition of "interesting". New phenomena could have unexpected, not-yet-theorized manifestations in the detector, and they would be missed using the current trigger algorithms. New phenomena could also be extremely rare and buried in uninteresting backgrounds, and require specialized data taking techniques. Moreover, disagreements of data with the SM prediction require dedicated investigation with the new dataset, and therefore dedicated trigger reconstruction and selection criteria. For example, the SM prediction that the weak force couples equally to all lepton types (Lepton Flavor Universality, or LFU) has recently been challenged by LHCb measurements [4]. This makes LFU and LFV (Lepton Flavor Violation, the prediction that the overall number of leptons of a given type does not change in interactions) among the most

[1] W. Riess, IBM Zurich, Nanoscience Colloquium at Lund University titled "The Future of Computing", 17/05/2018 [2] CERN, The LHC Experiments, CERN Document Server [3] Doglioni and Boveia, Dark Matter searches at colliders, Ann.Rev.Nucl.Part.Sci. 68 (2018) 429-459 [4] LHCb Collaboration, Test of lepton universality with $B^+ \to K^+\ell^+\ell$, Phys. Rev. Lett. 113, 151601 (2014). LHCb Collaboration, Test of lepton universality with $B^0 \to K^{*0}\ell^+\ell$, JHEP 08 (2017) 055.

interesting topics for the near-future of particle physics. Measurements of LFU and LFV are only one of the many examples that need RTA to be explored in detail with the upcoming LHC dataset. SMARTHEP will train a new generation of researchers with a strong expertise on RTA techniques and analysis of large datasets. The ESRs will **exploit innovative ML techniques and hybrid computing architectures to upgrade the trigger systems of LHC experiments for the next LHC data taking periods where the volume of data will increase tenfold**. The ESR in SMARTHEP will deploy the necessary RTA tools and use them to **perform measurements to verify the SM to a high precision and to search for new, unexpected physics phenomena that can shed light on the open questions of the fundamental components of matter**.

As detailed in Secs. 1.4.1 and 3.4, researchers in the SMARTHEP consortium are experts on the trigger systems of all four main LHC experiments. Within SMARTHEP, they will pool their expertise and complement it with that of computer scientists, entrepreneurs and professionals. With this consortium, SMARTHEP can **strengthen European innovation capacity, recognizing that RTA methods solve both academic research and industry challenges**, can be transferred from research to industry and vice versa leading to concrete deliverables.

SMARTHEP's ESR projects solve both research and industry challenges, as time-to-insight from raw data is a crucial benchmark for a competitive and effective use of the information. For this reason, advancement in terms of fast and efficient data analysis and decision-making are required by society and the commercial sector, and a selection of example industrial use cases sharing common issues with HEP has been chosen as integral part of the SMARTHEP research and exploitation program (see Sec. 2.3.2). RTA is a key point for improving transport and its safety, as well as for optimizing industrial processes. The information on the position of private or public transport vehicles and their possible routes are transmitted to a central analysis system, so that the traffic conditions around it are analyzed and a forecast is made. All of this needs to be done on a timescale that is short enough for a city to modify its traffic light system to avoid congestion, or to reroute vehicles to avoid delays. This requires both ML and hybrid computing architectures for algorithm training, similarly to complex problems such as tracking particles in a HEP detector. The detection of a driver's behavior (e.g. drowsiness) in fleets of cars should occur in milliseconds to preserve safety, and detection of malfunctioning components in industrial processes must trigger maintenance before the component breaks. Both those use cases require the optimization of advanced algorithms on limited computational resources, much like HEP's trigger systems. Another use case for RTA in industry is fraud detection, where spotting an anomalous transaction and deciding whether it is allowed or not needs to be made on a short timescale. This is a similar challenge to HEP triggers looking for unexpected phenomena. While ML algorithms trained to detect outliers are the current state of the art for this purpose, their black box approach raises interpretability and explainability challenges. This is one of the topics beyond the state-of-the-art that is tackled by SMARTHEP: knowledge-based approaches for the induction of explicit decision rules are useful both for increasing the accountability of fraud detection and for HEP trigger systems where every decision to discard an event is final and must be understood.

In summary, SMARTHEP trains researchers in ML, AI and data analysis and hybrid computing architectures, so they can enable RTA for the advancement of decision-making as well as monitoring and discoveries in both research and industry. The work and expertise developed within SMARTHEP will develop and commission upgraded trigger systems for LHC experiments for HEP measurements and searches, improve transport, industrial process and financial decision-making. [5]

Consortium composition

SMARTHEP spans ten different countries. CERN, the NIKHEF and CNRS research organizations, and universities SU, UH, UNIGE, TUDO, ULUND, UHEI form the core academic side of the Network, with experts in HEP and computer science. IBM and KKT as beneficiaries, and XIMANTIS, LIGHTBOX, POINT8, as partners, form a balanced counterpart to academia with a focus on the same technological challenges as HEP to gain insight into data in real-time. UO, OSU, INFN, USC, ULIEGE, RU and VU are high-profile associated academic partners whose expertise reinforces the network, that provide training and secondment to the ESRs, and award PhD degrees in case the beneficiaries are industries and research institutes[6]. All consortium members have successfully proven their capabilities for research, training, exploitation and dissemination as shown in Sec. 1.3, 3.4 and part B2.

Network goals and research objectives

The goal of SMARTHEP is to train ESRs in ML techniques and advanced data analysis, hybrid computing architectures, so they can advance real-time decision making and real-time monitoring of complex processes and discovery of unexpected phenomena.

[5] In line with the H2020 Societal Challenge for Smart, green and integrated transport. [6] In this proposal, we use "node" to indicate either beneficiary or partner. The *node responsible* is the *main scientific contact person* for the node. Administrators, albeit indispensable for the network functioning, are not named even if they are node responsibles in the case scientists are affiliated to two nodes.

These training goals define the four main research topics for SMARTHEP, each corresponding to a research Work Package (WP) detailed in Sec. 1.1.2, with research objectives and outcomes shown in each table in this Section that can be reached thanks to collaboration between industry and HEP. The first two topics, ML and hybrid architectures, provide the research tools so that RTA methods can used for the latter two topics, with concrete deliverables in physics and industry listed in Sec. 3.1.1.

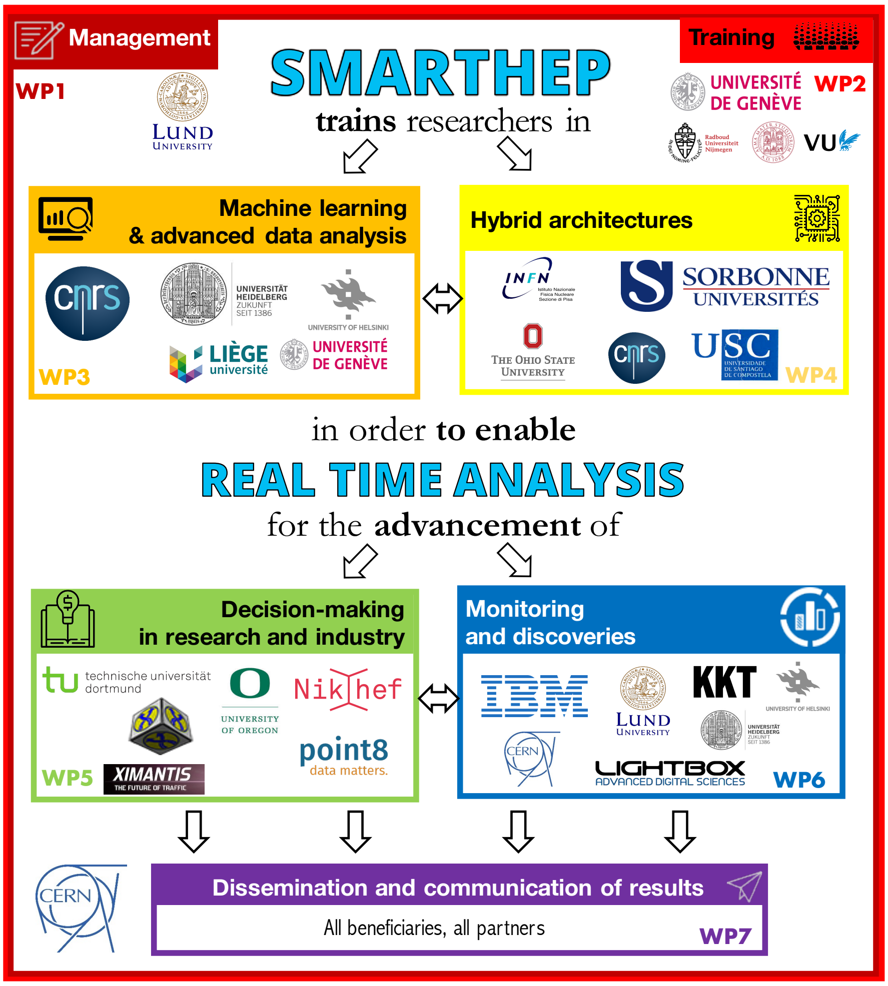

Figure 1: SMARTHEP network goals and implementation strategy, corresponding to WPs. The nodes with their main expertise in a given WP are represented in the respective WP box.

The first network goal is to train ESRs in the use of **ML and advanced data analysis** for use in in RTA techniques. ML techniques are by now ubiquitous in both HEP and industry, in order to optimally analyze raw, unstructured data. In SMARTHEP we will further extend their use in RTA and event reconstruction, paying particular attention to their understandability and reproducibility (see research topic 4). Meta-analysis (benchmarking) is also necessary to understand the algorithm performance prior to using them in RTA and to satisfy he requirements of resource-constrained environments.

Research Objective:	Outcomes:
design and use a variety of ML algorithms for RTA (GAN, RNN, Deep Learning)	toolkits to be used in HEP and commercially;
implement efficient real-time object and event reconstruction	software for HEP experiment triggers;
benchmark and optimize RTA algorithm performance	inter-sector benchmarking toolkits;

The second research topic focuses on the design of innovative **hybrid computing architecture** solutions in hardware/software, given that the complexity and rate of LHC data and beyond does not allow standard processors or data analysis techniques to be competitive[7]. An example of such a technology in HEP is the Fast TracKer in ATLAS, a hardware-based ATLAS Fast TracKer (FTK)[8], a unique hardware processor made by several custom electronic boards based on Field Programmable Gate Arrays (FPGA) to reconstruct the trajectory (track) of the charged particles that cross the inner part of the experiment. In industry, Graphic Processing Units (GPU) are used for computing parallel processes and training ML algorithms, and their use is being prototyped in HEP by members of SMARTHEP.

[7] e.g. J.P. Vlimant, Machine Learning in Charged Particle Tracking, Hammers and Nails ML&HEP Conference, 2017 [8] J Adelman et al., ATLAS FTK Challenge: Simulation of a Billion-fold Hardware Parallelism, Procedia Comput.Sci. 66 (2015) 540-545

Research Objective:	Outcomes:
use of FPGAs for fast hardware computations	trigger improvements for LHC experiments;
use of GPUs for massively parallel computations	significant speed-up for HEP and industrial ML and RTA techniques;
design parallelized and multithreaded algorithms	RTA software needed for the increase in LHC data and for more efficient financial transactions;

The third research topic concerns **real-time decision making in physics and industry**. The outcomes of the previous research topics will be applied to concrete physics cases and industrial challenges. These improvements in RTA benefit each experiment's physics program and increase the information quality extracted from data-rich environments in industry. There are many HEP drivers behind the trigger improvements that will be the SMARTHEP ESR's thesis topics: hints of physics beyond the SM in LFV/LFU, the mystery of 85% of the matter content of the universe that is not yet included in the SM, the in-depth study of the newly discovered Higgs boson and the study of collective particle behavior from states of matter created shortly after the Big Bang. Similar techniques used for trigger improvements will also help improve transport and fleet safety.

Research Objective:	Outcomes:
deploy improvements from real-time reconstruction and hybrid architectures in HEP trigger systems	measurements and searches with LHC Run-3 data;
use ML and hybrid architecture for RTA to improve transport and safety of fleets	inputs to traffic management of connected cities, extended transport apps and in-vehicle software;

The fourth research topic concerns efficient and accountable **monitoring** of complex processes and increase the sensitivity to physics **discoveries** through outlier detection and novel RTA techniques in data taking.

Firstly, industry and HEP share the challenge of not knowing in advance what kind of "interesting" events will appear, in the case of malfunctioning equipment, fraudulent financial transaction or potential not-yet-theorised new physics process. Outlier/novelty detection is a growing field, with dedicated ML algorithms[9]. In SMARTHEP, experts in the field and their supervised ESRs will apply those to physics and industry use cases, and extending them to methods where rules on what are inferred from a mixture of theory and known events to increase their accountability and understandability.

Secondly, the present HEP paradigm is to select in real-time a small number of interesting events, but keep the full detector data for later analysis. This paradigm depends on the assumption that interesting events are very rare compared to the SM backgrounds. It is however becoming increasingly clear that crucial HEP research questions can only be answered by either precise measurements of the abundant SM processes themselves, or by searching for new processes which closely mimic these SM backgrounds [10]. To be sensitive to these processes, higher-level physics insight must be obtained in real-time. This way, only a small subset of this higher-level information can be stored for each event, greatly reducing its size. This technique is referred in the following as *fully RTA-based data taking*, it has been pioneered by SMARTHEP members and published in one of the highest-impact journals in the field [11].

SMARTHEP will improve and extend this technique beyond the current state-of-the-art towards the upgrades, in LHCb where most of the data will be recorded this way, in ALICE where it will be implemented for the first time, and in ATLAS and CMS where it will be applied to a larger number of searches.

Research Objective:	Outcomes:
implement outlier detection techniques in physics and industry	potential for not-yet-theorized discoveries at the LHC, improved fraud detection algorithms and IoT-ready industrial predictive maintenance;
study and design rule-induction algorithms as complement to ML to improve accountability and understandability	RTA for fraud detection and HEP triggers;
extend RTA data-taking techniques in HEP trigger systems where raw data is discarded	increased sensitivity to new physics processes in LFV/LFU, Higgs boson measurements and dark matter searches

B1.1.2 Research methodology and approach

In SMARTHEP, cutting-edge techniques in ML and hybrid architectures will be assessed and developed by the ESRs. They will be applied to specific use cases of RTA that advance HEP, industry and society, and the results exploited, disseminated and communicated (Sec. 2.3). This modus operandi is at the heart of SMARTHEP and ensures the efficiency of its implementation.

SMARTHEP defines WPs for the management of the consortium (WP1), for training (WP2), and for outreach and dissemination, and four research WPs, corresponding to each of the research topics in the previous section. The WPs are summarized in the Table 1.1. Fig 1. gives a schematic view of the four research WPs and the main actors within the consortium.

[9] R. Domingues et al, http://doi.org/10.1016/j.patcog.2017.09.037, Pattern Recognition 74 (2018) [10] See e.g. Gligorov and Fitzpatrick, Anatomy of an upgrade event in the upgrade era, and implications for the LHCb trigger, CDS Document Server and ALICE Collaboration including Christiansen, Technical Design Report, CDS Document Server [11] CMS Collaboration including Pierini, Search for narrow resonances in dijet final states at $\sqrt{s} = 8$ TeV with the novel CMS technique of data scouting, Phys. Rev. Lett. 117 (2016) 031802. ATLAS Collaboration including Boveia, Starovoitov, Dunford and Strom, Search for low-mass dijet resonances using trigger-level jets with the ATLAS detector in pp collisions at sqrt(s)=13 TeV, Phys. Rev. Lett. 121 (2018) 081801, LHCb collaboration with Borsato as lead author https://arxiv.org/abs/1710.02867, Phys.Rev.Lett. 120 (2016) 061801.

WP No.	WP Title	Lead Beneficiary No. / responsible	Start	End	Activity Type	Lead (short)	ESRs
WP1	Management	1, Doglioni	1	48	Management	ULUND	-
WP2	Training	8, Sfyrla	5	42	Recruitment, training	UNIGE	All
WP3	ML, AI and data analysis	2, Gligorov	1	48	Research	CNRS	1-5, 7, 9-14
WP4	Hybrid architectures	9, Lacassagne	1	48	Research	SU	4-9, 14
WP5	Real-time decision making	3, Albrecht/Sopasakis	1	48	Research	TUDO	5, 7, 8, 10-14
WP6	Monitoring & discoveries	10, de Sainte Marie/Pierini	1	48	Research	IBM	1-4, 6, 7, 14, 15
WP7	Dissemination & communication of results	6, Petersen	1	48	Dissemination, outreach	CERN	All

Table 1.1: Work Package (WP) List

In **WP3**, we make full use of advanced **Machine Learning, AI and data analysis** techniques in RTA. This is an emerging field of research that enables real time decision-making, monitoring and discoveries in industry and society. The ESRs involved in this WP will use ML to speed up the reconstruction of raw detector information and use it for decision making (ESRs 1,7,13-14), develop, optimize and benchmark algorithms for use in resource-constrained environments (ESRs 4,6,9-11) and detect outliers (ESRs 2,3). In **WP4**, we will develop algorithm implementations which efficiently use modern computing hardware that is increasingly both highly parallel (ESR8) and heterogeneous, with CPU-based processing farms complemented by GPU or FPGA co-processors (ESRs 4-7,9,14). This kind of **hybrid architectures** enables faster data processing, and their design requires highly-specialized training that the consortium including computer scientists is best placed to deliver. In **WP5**, we will employ the techniques and tools developed in WP1 and WP2 to enable advanced RTA techniques for **decision-making**. In HEP, we will develop and commission the upgraded triggers of the main LHC collaborations and use them for physics analysis (ESRs 5, 7, 8, 10-14), while in industry we will improve traffic and fleet safety (ESRs 4,5,10,12). Within **WP6**, we will focus on **real time monitoring of complex processes and preparations for unexpected discoveries**. This feeds back to the decision-making WP5 through novelty/anomaly detection algorithms that are understandable and reproducible, in HEP (ESR 2,3), finance (ESR 3) and predictive maintenance (ESR 7). We will perform physics analysis that is fully based on RTA without the need to save raw data (ESRs 1,5,12,14,15). All results will be **disseminated and communicated** to a variety of audiences within **WP7**, paying particular attention so that publications, software and documentation are accessible as legacy of the action according to the FAIR principles. HEP experiments are by their very nature built around **training**, as around a third of all LHC collaborations consist of PhD students. The training WP (**WP2**) encompasses WP3-7 and combines this collaborative, training-based research culture, with a focus on cutting-edge technologies.

Throughout this ETN, RTA will provide the common language linking these challenges and disciplines, in research, academia and industry, and will enable each of the WPs to benefit from the progress made in the others. In particular, WP3 and WP4 will provide the techniques to enable advancements in WP4 and WP5. For this reason, most ESRs will start their research focusing on either ML or hybrid architecture then proceed to apply the knowledge and technique to the decision-making or monitoring and discoveries, as shown in Fig. 4.

B1.1.3 Originality and innovative aspects of the research program

The ESRs in the SMARTHEP network will undertake a timely and ambitious research program on RTA, benefitting from the experience not only of their direct supervisors, but of the entire network, to reach their objectives. Some of the original and innovative aspects of the projects in the network are listed below.

1. Researchers from SMARTHEP process large datasets using RTA techniques in a cost-effective way. A part of the novelty of this proposal is the volume of data which will be processed in real-time, comparable to the largest commercial tasks, but on a much smaller budget. One can compare Facebook and e.g. the LHCb collaboration. The former spends hundreds of millions of dollars a year on computing[12] while the latter only spends order 10M dollars a year on computing [13]. To maintain a cost-effective data taking strategy, LHC experiments need a more systematic application of RTA (especially fully RTA-based analysis that discards bulky raw data), machine learning and hybrid architectures for HEP. As RTA is essential in industry as well as in HEP, the research environment can benefit from cross-pollination in terms of research topics, and from a generation of ESRs with experience in industrial grade algorithms and tools.

2. The SMARTHEP research program could lead to breakthroughs in our understanding of nature Only fully-real time data acquisition techniques and trigger improvements enable full exploitation of the LHC dataset. The HEP research topics chosen to drive conceptual developments within SMARTHEP have the potential to lead to the discovery of new physics beyond the Standard Model. Examples include studies of lepton flavor and its conservation and universality in different final states

[12] Facebook, The Facebook Data Center FAQ, 2010. [13] Private communication, Prof. Peter Clarke, University of Edinburgh.

and experiments, dark matter mediators and new light particles, precise probes of the Higgs boson and of heavy ion physics, as well as not-yet-theorized phenomena that would not be possible to record using the current state-of-the-art triggers.
3. ESRs in SMARTHEP deploy and disseminate their research at a unique time for particle physics. As highlighted by the HEP Software Foundation Whitepaper[4], the period 2019-2023 is ideal for this R&D in HEP, as it is a time of transition between LHC data taking periods that will be necessary to prepare for an upgrade of the LHC accelerator where the amount of data delivered will make RTA techniques the key to pursue the physics programs of the four main LHC experiments. The systematic optimization of HEP experiments by SMARTHEP will boost the performance of the current and planned upgrades of the CERN based accelerator experiments. Furthermore, the developed toolkits will be advertised at international conferences and thus the developed methods will shape the online event selection of all future HEP experiments.
4. SMARTHEP researchers further the RTA paradigm in both HEP and industry. The close links of the research institutes of the consortium with the industry partners means that the ESRs will directly drive the development of novel industrial products, while also bringing professional methods of data mining that are exercised in large companies into the academic environment. Most modern methods applied in research can in this way be transferred to industry applications. By exposing industry-grade methods to the volume and complexity of HEP data, we will stimulate their development in a complementary way for the benefit of industry. In particular, the ESRs will be trained in topics matching the emerging trend of real-time big data analytics[14] by invited speakers from large enterprises using these methods.

B1.2 Quality and innovative aspects of the training programme

The training program of SMARTHEP has three main **objectives**, taking the Salzburg Principles[15] as a guidance:

- **Objective A**: provide ESRs with knowledge and training to conduct original research during and beyond their PhD studies;
- **Objective B**: provide solid bases in a broad spectrum of topics related to their field of research, extending beyond their dedicated research and including soft skills;
- **Objective C**: provide up-to-date and career-related training through interactions with multiple collaborators within academia as well as the industry, to meet the needs of a broad employment market.

Objectives A and B will be fulfilled by schools, workshops and events organized within the network; ESRs will profit from high quality lectures delivered by experts within the institutes associated to SMARTHEP, complemented by doctoral training at local nodes. Objective C will be fulfilled by mentoring by senior scientists and special network events where the industrial collaborators will provide seminars and training on commercial applications of SMARTHEP's objectives. All these objectives will be further addressed via the natural collaboration between the ESRs and the various institutes of the network, as well as the planned secondments in academia and industry. This is a more flexible and diverse structure compared to standard PhD studies and will allow the ESRs to develop an individual and unique research mindset within an inclusive environment, so that they can act as proactive participants in furthering HEP and industry goals through RTA techniques.

B1.2.1 Overview and content structure of the training

Training is at the heart of all activities within SMARTHEP, with a dedicated Work Package, WP2. All ESRs that will be recruited will obtain a PhD, have a Personal Career Development Plan (PCDP) outlining the training through research, secondment and lectures at the local and host nodes that includes transferrable skills training, and they will be able to attend complementary network-wide training. Their training will be accounted through a credit system designed for this ETN.

The recruiting deliverables, the composition of the Supervisory Committee (SC) for each ESR and the expertise and qualification of supervisors is shown in Tab. 1.2.

ESR	Recruiting node	PhD-awarding node	Planned start	Duration	Title
ESR1	UH	UH	8	36	ML and RTA for Higgs boson measurements and industry
SC	**Main supervisor: Voutilainen** (UH), [2]. Assistant prof. Expertise: RTA, SM and Higgs measurements, with positions of responsibility in CMS. Grant from the Finnish Academy of Science, prizes for research during PhD and postdoc. **Second supervisor: Pierini** (CERN) [6]. Staff scientist. Expertise: RTA, ML for HEP. ERC CoG for ML at the LHC. **Industrial supervisor: Taccari** (KKT). Data scientist. Expertise: mathematical optimization and ML. **Additional supervision: Kirschenmann** (UH). Postdoc. Expertise: LHC jet physics and searches, with positions of responsibility in CMS. **Senior mentor: Eerola** (CERN). Professor (HEP), Vice Rector of the University of Helsinki.				
ESR2	IBM	ULUND	8	36	Novelty detection for industry and ATLAS searches
SC	**Main supervisors: Jullié** (IBM), Software Engineer. Expertise: modelling optimization problems, ML, anomaly detection. **Doglioni** (ULUND), [6], Assistant prof. Expertise: jet, trigger and Dark Sectors expert in ATLAS, LHC Dark Matter WG organizer and HSF trigger and reconstruction WG convenor, ERC StG 2015 on RTA for ATLAS, national grants. **Second supervisor: Boveia** (OSU). Expertise: Dark Sectors, trigger and track trigger ATLAS, responsible for first fully RTA-based search in ATLAS. **Additional supervision: Shaw** (IBM). Expertise: constraint programming, optimization modelling and local search. **Senior mentor: Akesson** (ULUND). Professor. Expert in searches at ATLAS. Ex-CERN Council president.				

Table 1.2: Recruitment deliverables per beneficiary, Supervisory Committee (SC) with number of supervised PhD students and postdocs, and expertise/qualification of supervisors. Continues on next page.

[14] A comprehensive guide to real-time big data analytics. [15] Salzburg II recommendations webpage

ESR	Recruiting node	PhD-awarding node	Planned start	Duration	Title
ESR3	IBM	SU	8	36	Real-time rule induction in fraud detection and HEP
SC	**Main supervisors: Feillet** (IBM), Software architect (Decision Lab). Product engineering of mission critical decision automation systems, fraud detection, integration of ML and rule induction. Experience in supervising PhD thesis with LIP6, board member of Association Francaise pour l'Intelligence Artificielle (industrial/academic). **Lacassagne** (SU), [15], Professor. Hybrid architectures expert, leader of LIP6 Hardware and Software for Embedded Systems team. **Second supervisor: Gligorov** (CNRS) [3]. Staff scientist. Expertise: ML, LFV/LFU, trigger, past LHCb Deputy Physics Coordinator, Real-time analysis project coordinator. ERC CoG 2016 on RTA and LFV/LFU in LHCb. **Additional supervision: Louppe** (ULIEGE) [3]. Assistant prof. Expertise: AI and deep learning, Analysis Consultant Expert with ATLAS.				
ESR4	KKT	UNIBO	8	36	RTA through computer vision on dashcams
SC	**Main supervisors: Sambo** (KKT). Senior Data Scientist. Expertise: AI, ML, connected vehicles. **Di Stefano** (UNIBO) [11]. Professor. Expertise: computer architecture and vision, image processing. **Second supervisor: Lacassagne** (SU) [15]. Professor. See ESR3. **Additional supervision: Salti** (UNIBO). Expertise: computing engineering, computer vision. **Doglioni** (ULUND). See ESR2.				
ESR5	UHEI	UHEI	8	36	Real-time noise reduction new physics searches
SC	**Main supervisors: Dunford** (UHEI), [8], Ausserplanmäßig. and Young Researcher group leader. Expertise: Dark Matter, SM measurements, detector and trigger expert in ATLAS, editor of first trigger-level analysis paper in ATLAS. **Starovoitov** (UHEI), [7], Postdoc. Expert in ATLAS detector, trigger and data analysis, SM measurements and searches, receiver of several international fellowships. **Second supervisor: Strom** (UO) [6]. Professor. Expert in trigger and data acquisition in ATLAS, responsible of all aspect of trigger and data acquisition (2017-2018) and FTK (2018-). **Industrial supervisor: Sopasakis** (XIMANTIS). CEO and ULUND associate prof. Mathematics and ML/AI expert, spin-off and start-up experience. **Additional supervision: Boveia, Doglioni** (OSU, ULUND). See ESR2. **Senior mentor: Hansmann-Menzemer** (UHEI). Professor. See ESR14.				
ESR6	CNRS	SU	8	36	Real-time trajectory reconstruction in ATLAS
SC	**Main supervisors: Crescioli** (CNRS and SU) [3], see ESR8. **Malaescu**(CNRS and SU), [4], Assistant prof. Expertise: jet measurements and searches, ATLAS SM group convener, past Statistics Forum convener. **Second supervisor: Roda** (INFN) [14]. Expert in searches beyond the SM, ATLAS Pisa group leader, coordination roles in ATLAS. **Industrial supervisor: Sambo** (KKT). Senior Data Scientist. See ESR4. **Additional supervision: Annovi** (INFN). Expertise: tracking and track triggers for ATLAS **Senior mentor: Calderini** (CNRS), Team leader, AIDA2020 management. Expertise: detectors, SM in ATLAS				
ESR7	UNIGE	UNIGE	8	36	ML pattern recognition for exotic physics and industry
SC	**Main supervisor: Sfyrla** (UNIGE), [2]. Assistant prof. (HEP). Expertise: SUSY DM and trigger expert, responsibility positions in ATLAS in trigger and HL-LHC upgrade. Grant from the Swiss National Foundation and Boninchi Foundation. **Second supervisor: Martinez-Santos** (USC) [6]. Assistant prof. Expertise: rare decays, trigger and detectors. ERC StG for BSM searches. **Industrial supervisor: Catastini** (LIGHTBOX). Quantitative Analyst. Expertise: analysis of financial markets, automated digital advertising trading. **Supervising postdoc: Schramm** (UNIGE). Expertise: ML and new physics, convener of LHC interexperiment ML group until 2018, Banting Fellow. **Senior mentor: Iacobucci** (UNIGE). Director of the DPNC at UNIGE.				
ESR8	CERN	UNIGE	8	36	Efficient RTA in ATLAS using multi-threaded processing
SC	**Main supervisors: Petersen**(CERN), [2]. Staff scientist. SUSY DM and trigger expert, past ATLAS trigger convener, current upgrade physics convener; **Sfyrla** (UNIGE), see ESR7 **Second supervisor: Crescioli** (CNRS) [3]. Research engineer. FTK and FPGA expert, technical coordinator of many R&D projects; **Industrial supervisor: Catastini** (LIGHTBOX), see ESR7. **Senior mentor: Monica Pepe-Altarelli** (CERN). Staff scientist. Vice-President Elected at Large of the Executive Council of IUPAP				
ESR9	SU	SU	8	36	RTA on heterogeneous computing architectures
SC	**Main supervisor: Lacassagne** (SU), [15], Professor. See ESR3. **Second supervisor: Petersen**, Staff scientist. see ESR8 **Industrial supervisor: Borri** (LIGHTBOX). Staff. Expertise in quantitative analysis and automated financial market trading expert. **Additional supervision: Couturier** (CERN). Expertise: software architect for various companies, now core team of LHCb software framework				
ESR10	NIKHEF	RU	8	36	Optimization of RTA resources and ATLAS LFV search
SC	**Main supervisor: Igonkina** (NIKHEF and RU), [14]. Professor. Expert in LFV/LFU and triggering in ATLAS with positions of responsibility, receiver of several national grants. **Second supervisor: Strom** (UO) [6]. Professor. See ESR5 **Industrial supervisor: Sopasakis** (XIMANTIS). CEO and ULUND associate prof. See ESR5. **Additional supervision: Petersen, Couturier** (CERN). Expertise: See ESR8 and ESR9.				
ESR11	NIKHEF	VU	8	36	Optimization of RTA resources and LHCb LFV search
SC	**Main supervisor: Raven** (NIKHEF and VU), [14]. Professor. expert in LFV/LFU and triggering in LHCb with positions of responsibility, receiver of several national grants. **Second supervisor: Albrecht** (TUDO) [13]. Emmy Noether group leader and lecturer. See ESR12. **Industrial supervisor: Brambach** (POINT8). Staff. Expertise: data science, sales and project management, PhD in LHCb. **Additional supervision: Petersen, Couturier** CERN. See ESR8 and ESR9.				
ESR12	TUDO	TUDO	8	36	Real-time ML for LFV in unflavoured meson decays
SC	**Main supervisor: Albrecht**(TUDO), [8]. Emmy Noether group leader and lecturer. Expertise: LFV/LFU, tracking, ML and trigger expert, deputy Physics Coordinator of LHCb. ERC StG 2016 on LFV/LFU. **Second supervisor: Martinez-Santos** (USC), [6]. Assistant prof. See ESR7. **Industrial supervisor: Sopasakis** (XIMANTIS). CEO and ULUND associate prof. See ESR5. **Supervising postdoc: Matev** (CERN). Expertise: software and trigger design and maintenance. **Senior mentor: Spaan** (TUDO). Head of experimental physics 5, team leader for the LHCb experiment, project leader of CS research area SFB876.				

Table 1.2: Recruitment deliverables per beneficiary, Supervisory Committee (SC) with number of supervised PhD students and postdocs, and expertise/qualification of supervisors. Continued from previous page, continues on next page.

ESR	Recruiting node	PhD-awarding node	Planned start	Duration	Title
ESR13	TUDO	TUDO	8	36	Global event triggering in LHCb
SC	**Main supervisor: Albrecht**(TUDO), [13]. Assistant prof. See ESR12. **Second supervisor: Raven** (NIKHEF and VU) [14]. Professor. See ESR11. **Industrial supervisor: Dungs** (POINT8). Staff. Expertise: transitioned from LHCb trigger group to data science (via Google). **Senior mentor: Spaan** (TUDO), see ESR12.				
ESR14	UHEI	UHEI	8	36	RTA to search for Dark Photons in LHCb
SC	**Main supervisor: Hansmann-Menzemer** (UHEI), [15]. Professor. Expertise: tracking algorithms and software in LHCb. Co-spokesperson of research training "Particle Physics beyond the SM" (DFG). EPS Young Researcher award (2007), ERC StG 2010 **Second supervisor: Albrecht** (TUDO) [13]. Emmy Noether group leader and lecturer. See ESR12. **Industrial supervisor: Dungs** (POINT8). See ESR13. **Additional supervision: Borsato** (UHEI). Expertise: Dark sectors, trigger and software for LHCb.**Martinez-Santos** (USC). Assistant prof. See ESR12.				
ESR15	ULUND	ULUND	8	36	Real-time calibration and analysis of the ALICE Time Projection Chamber
SC	**Main supervisor: Christiansen** (ULUND), [17], Professor. Expert in Time Projection Chambers and reconstruction of their data. Expertise in data analysis in ALICE. Receiver of several national grants. **Second supervisor: Shahoyan** (CERN) [2]. Staff scientist (LD). Main developer of ALICE real-time reconstruction project. **Industrial supervisor: Sopasakis** (XIMANTIS). CEO and ULUND associate prof. See ESR5.				

Total months: 540

Table 1.2: Recruitment deliverables per beneficiary, Supervisory Committee (SC) with number of supervised PhD students and postdocs, and expertise/qualification of supervisors. Continued from previous page.

Recruitment deliverables per participant and awarding of PhD degrees. ESRs will be recruited by month 8, complete a total of 36 months of research and training, and be awarded a PhD degree. ESRs at non-academic beneficiaries and ESRs at international or national laboratories will be awarded PhDs by universities within the network and have supervisors in that university as well. Each ESR will have a Supervisory Committee (SC) composed of participants from the local and secondment academic and industrial nodes. For Finland, Netherlands and Sweden which mandate a four-year PhD, all beneficiaries will provide support for the ESRs to complete their PhD thesis.

Personal Career Development Plan. Two months after the start of their projects, ESRs and their SC will present a **Personal Career Development Plan** (PCDP) so that the **core, advanced and transferrable skills** to be acquired, as well as the milestones for each of the ESR projects, can be agreed between student and supervisor and consortium, planned and monitored throughout the course of the program, taking into account the existing resources both at the ESR node and at foreseen secondments. The PCDP will be reviewed by the local node coordinator, who will bring it to the Supervisory Board (SB) for approval. The main supervisor will be available throughout the course of the PhD, and meet with the ESR on a weekly basis and with an open-door policy. The local node coordinator will also review the progress described in the PCDP at least every six months (e.g. during staff appraisal meetings) and bring a short report to the consortium. The PCDP[16] will include the requirements, milestones and goals within the schedule of the doctoral program (including secondments), the local and network-wide courses and schools to be attended, and a list of dissemination, communication and outreach activities.

Network-wide events: schools, yearly meetings and schools. Network-wide schools, conferences and events shown in the table below will be organized by SMARTHEP beneficiaries and partners as part of the training and dissemination program and its preparation. We expect ESRs to attend network events in person, but wherever possible will make network school, conferences and events available as Webinars using the Vidyo technology provided by CERN, to allow all SMARTHEP ESRs and PIs to attend if family/personal commitments would otherwise prevent it. A permanent record of the lectures will be available as proceedings, and in some cases video recordings, as SMARTHEP has the ambition to make the training program available beyond this Action and continue organizing successful schools. The table below summarizes all events included within the network, with compulsory schools marked in bold so that students attend a yearly meeting and a school each year, dedicating sufficient time to local training and research. The hosts and lecturers for these events schools have been identified within the network based on their expertise, see Sec. B1.4.1.

Main training events and conferences	Credits	Lead (support) institution	Action month	Notes
1. **Kick-off meeting**	-	ULUND	2	-
The kick-off meeting will be dedicated to organizing the project management, signing the consortium agreement, and monitoring the ESR recruitment.				
2. **Yearly conferences**	2	ULUND (NIKHEF, IBM)	9,23,36	-
SMARTHEP will hold **yearly in-person network-internal conferences**, with a duration of 4 or 5 days. On the first two days, the ESRs will gain experience in presenting their research by giving presentations to the other SMARTHEP members and participating in a poster session. Days 3–5 will be dedicated to lectures on research, technical and transferable skills, open discussions of synergies between industrial participants, as well as communication and dissemination activities tailored to local circumstances (see Sec. 2.4.1). The first yearly meeting on month 8 will serve as introductory school on HEP and RTA, as it will be hosted shortly after all ESRs have been recruited. All yearly meetings will include a dedicated half-day of lectures on non-academic training. Within the Yearly Meetings, there will be time devoted to management activities such as the Executive Board, ESR Board and Supervisory Board meetings (see Sec. 3.2.1), that also train ESRs in scientific collaboration and governance.				

Table 1.3: Network events and schools. Continues on next page.

[16] The SMARTHEP PCDP will follow a common template based on that from the MSCA website to be coherent between the ESRs.

Main training events and conferences		Credits	Lead (support) institution	Action month	Notes
5.	**Non-academic training workshop**	1.5	ULUND	10	Network event joint w/INSIGHTS ETN
This workshop will follow the first yearly meeting and be held in either the University of Oslo (UiO) or ULUND. From preliminary discussions, the coordinator of the INSIGHTS ETN (where ULUND is a beneficiary and C. Doglioni is a diversity and inclusion officer) and node responsible for this training have agreed to make it a joint event if SMARTHEP is funded. In this workshop, the ESRs will receive non-academic training by experts at ULUND and UiO, as well as INSIGHTS' partner KPMG. The topics of the lectures provided by SMARTHEP will be diversity and inclusion, team-work, research integrity and sustainable research, as an early complement to local training on transferrable and soft skills so that the ESRs can start their work in a positive environment. The joint organization of ULUND and UiO, both within the INSIGHTS ETN, will allow ESRs from two different ITNs to meet, exchange ideas and experiences, and broaden their network.					
4.	**Physics and machine learning school**	3	UNIGE	15	New network event
This school will provide all ESRs with more advanced courses on the physics topics tackled in the network, a few months after the ESRs have started their projects. It will also provide an introduction to how to design a physics analysis, as well as to machine learning concepts and their connections to HEP and industrial applications.					
5.	International School of Trigger And Data AcQuisition (ISOTDAQ)	3	UO	18 or 30	-
ISOTDAQ is a yearly school dedicated to triggering and acquiring data for physics experiments with lectures and hands-on exercises in equal proportions. A lecture about RTA by UO members of the SMARTHEP network will be added to the program if funded. ESRs will be encouraged to follow one of the two editions of the school during their PhD.					
5.	Machine Learning for HEP school (ML-HEP)	3	CERN	11 or 22	-
MLHEP is a school on cutting-edge machine learning techniques featuring dedicated trigger lectures, whose main organizer (Ustyuzhanin) is attached to CERN as a LHCb member. ESRs will be suggested to follow one of the two editions of the school during their PhD, and the nodes in SMARTHEP will be encouraged to place a bid to host the school as ULUND did in 2016.					
6.	Basic FPGA course, FPGA boot-camp	3	OSU, CNRS	26, 27	New network event
This school will include lectures on technologies and architecture and hands-on exercises on triggering applications. This school will consist of introductory courses given at CERN, and a follow-up bootcamp in Paris, where practical problems are solved for ESRs specializing in this topic. Lectures will be given from researchers in OSU, CNRS with the help of INFN, all leading institutes in research and development on hardware track triggers.					
7.	GPU and hybrid architectures school	3	USC (SU)	29	New network event
In this school, the ESRs will learn how to compare architectures and practical solutions, and programming on different platforms (e.g. GPU programming). The ESRs will learn what is available and planned on the market, as well as ways to best evaluate the chosen hardware solution for the software they are developing.					
8.	**Industry & career development school**	3	TUDO (XIMANTIS)	38	New network event
This school will include lectures and workshops in collaboration with industry, and it is optimally placed towards the end of the ESR's PhD training in month 44. This school is dedicated to an in depth study of strategies for intellectual property rights (IPR), commercializing research output, presenting research results to policy-makers, and knowledge transfer. One day of this school will be dedicated to group work on case studies prepared by the industrial beneficiaries and partners. This school will include experiences and Q&A sessions with the CEOs and founders of the companies within SMARTHEP.					
9.	**Final public conference**	2	CNRS	42	Yes
We will hold a five-day conference which will showcase the work of the network to the wider scientific community. As opposed to the yearly meetings, the conference will not feature management meetings, and will be dedicated to presentations on SMARTHEP research. Invited topical presentations on state-of-the art developments within the HEP and industry. Each of the conferences will dedicate between half and one day to dissemination activities by SMARTHEP members. The conference will have two additional days reserved for ESRs' presentations and public lectures showcasing the work done during their PhD program with SMARTHEP. ESRs will also take active part in the organization of this conference, to add to their transferrable skillset.					
10.	**Closing meeting**	-	UHEI	48	
In the closing meeting, the network will take stock of the experience of the ETN and plan the next steps. The closing meeting is beyond the PhD period for most of the ESRs, who will be invited as network alumni to give feedback on the ETN experience. The PIs and the ESR alumni will give public lectures.					

Table 1.3: Network events and schools.

All network-organized schools will be also open to the local students of the beneficiary and partners organizing the school and advertised through the HEP Software Foundation Training group, as part of enhancing the overall training program of the institutions involved in SMARTHEP and broaden participation to the network's activities. The HEP Software Foundation will support training activities within SMARTHEP, demonstrated in the Letter of Support at the end of this document. External students will be accepted if the optimal capacity of the school is not reached by the participating ESRs alone, and a fee will be charged to non-network participants if the node incurs external expenses due to their participation.
Local training. Training events organized specifically for SMARTHEP are complemented by local training provided by beneficiary nodes. Students from the network will be able to participate in this training when located at the node. All nodes include a range of graduate level courses in languages, career management, presentational skills, diversity and inclusion, as well as pedagogical courses in teaching and learning that the ESRs will be encouraged to follow to obtain the necessary amount of credits. Partners will also contribute with individual supervision and local training for employees while students are seconded at their premises. Links to the training courses for each node are detailed in their description, while below we list examples of graduate schools and training programs within the nodes that the ESRs will be associated to.

Beneficiary	Available graduate school or program
ULUND	COMPUTE graduate school in advanced computing techniques for research
TUDO	Bi-yearly graduate schools in the special research area *Providing Information by Resource-Constrained Data Analysis* (SFB876)
UHEI	HighRR research training group on detector development.
UH	HEP graduate school (40 ECTS points, of which 10 of transferrable skills)
CERN	Academic training program including transferrable skills.
NIKHEF	Research school in sub-atomic physics.
UNIGE	HEP doctoral school in theory and phenomenology.
SU and CNRS	Dedicated doctoral school for sciences, including courses on collaborative software and GPU.
IBM	Summer schools from the Association Francaise pour l'Intelligence Artificielle (AFIA).
KKT	Dedicated courses in computer science and image processing, in association with CVLab at UNIBO

As SMARTHEP also intends to prepare ESRs to teach others, specific credits will also be assigned to teaching skills. This includes taking pedagogical courses, as well as supervising Bachelor's and Master's students at the local node.

SMARTHEP credits system:

Table 1.4: Example SMARTHEP doctoral program

Type of training	No. of credits
Training through research	**135**
At host	75
Through secondment	60
Training through lectures, courses and dissemination	**45**
Technical and Research Training	30
Transferable Skills Training	15
of which towards teaching and dissemination	5
Total	180

To ensure that such a diverse training program is coherent and recognized across the network, we have designed a SMARTHEP credit system according to the ECTS standard. Each ESR will complete 180 SMARTHEP credits, as shown for an example ESR with two secondments in Table ??. As PhDs in all participating institutions are awarded based on local regulations, SMARTHEP credits also ensure that ESRs receive the appropriate training. Each network-wide event below includes the amount of assigned SMARTHEP credits.

We have assigned 1 SMARTHEP credit per each 1/2 day of lectures. All students will have to explicitly include 15 credits of transferrable skills training within their PCDP. By attending the yearly meetings of the network and the final conference, around 1/2 of the SMARTHEP credits that they will need to complete in both "Technical and Research" and "Transferable Skills" categories, as presented at the beginning of the section, will be provided in network-wide events. ESRs will have freedom how to complete the rest of required credits through the local resources. Finally, ESRs will be required to present their work to at least one conference outside the network in their area of expertise. A list of conferences of interest for SMARTHEP topics is given in Sec. 2.3.1. The amount of credits awarded will depend on the targeted conference. The attribution of ECTS credits will require institutes to explicitly include the network events in their course plan, but the conversion from SMARTHEP to ECTS credits will be justified and straightforward.

B1.2.2 Role of non-academic sector in the training program

SMARTHEP dynamically and directly addresses the challenges of the academic, industrial, and entrepreneurial sectors. This enriches all sectors through the transfer of best practices, knowledge, and expertise. For this reason SMARTHEP includes a comprehensive program of non-academic training, with hands-on experience in solving practical problems during secondments at industrial partners, discussed further in Sec. 1.4.3. The non–academic partners of SMARTHEP will have the following roles in the training of the ESRs:

- **Training through research: secondments**. One of SMARTHEP's most important objectives is to increase the exposure of the students to the private sector, solving practical problems with RTA and creating commercial value. Most ESRs will have secondments at private companies relevant for their tools and research topics. These will place a particular emphasis on common methods between the commercial applications which the non–academic partners specialize in, and the academic goals of the ESR projects. The connections are emphasized in the ESR project descriptions (Sec. 3.1.5).
- **Training through mentoring and supervision:** Industrial secondment supervisors are co-supervisors for the overall PhD project. Only three out of 15 ESRs don't have an industrial secondment. To ensure that they are also exposed to non-academic environment they will be assigned a non–academic mentor throughout their PhD.
- **Training through lectures:** All industrial beneficiaries and partners have also agreed to provide lectures and courses on transferrable skills and on their experience in occasion of the yearly meetings, in the non-academic training workshop and in the industry school. In particular, participants from LIGHTBOX, POINT8 and KKT have an academic background, participants from UNIBO transitioned from industry back to academia, and participants from XIMANTIS have a start-up in parallel to their academic position. This places the consortium in the best position to prepare the ESRs for a fluid career in academia and industry.

B1.3 Quality of the supervision

B1.3.1 Qualifications and supervision experience of supervisors

Table 1.2 summarizes the qualifications and supervision experience of the supervisors in SMARTHEP. Each ESR will have a SC that includes the main experts in in RTA, machine learning and hybrid architectures from all LHC collaborations. They have an excellent track record demonstrated by the European and national grants received by the PIs (5 ongoing ERC H2020 grants and numerous national grants). Each SC includes an industrial supervisor related to the topic developed in the PhD project that will also provide career-related mentoring.

Many of the SMARTHEP academic participants are young yet successful scientists with long-term academic positions, who dedicate a large fraction of their time to research and direct supervision of students. This guarantees that, as specified in the European Charter for Researchers, they will have the time, knowledge, experience and commitment to offer support and feedback so to guarantee the supervision of excellent research. To complement the supervision, ESRs supervised by young professors also have a senior professor/researcher from their local node included in their SC. The addition of research staff the

ESR SC guarantees that the ESRs receive day-to-day hands-on local or remote supervision, and allows SMARTHEP postdocs to gain supervision experience necessary towards their future career to positions in academia and industry.

All members of the ESR SC in Table 1.2 have **proven experience in supervising PhD theses**, as shown by the numbers of supervised early stage researchers from the PhD level onwards. Most of the researchers in the network, including the postdocs, have also held positions of responsibility within their large international collaborations. Within these roles, they have trained numerous students who performed research tasks under their supervision, and they have managed large groups of researches to successfully operate complex detector systems. Supervisors in industrial nodes have supervision expertise from **trainees to working in their company**, and are therefore well placed to deliver innovative cross-disciplinary training within SMARTHEP.

The ESRs also will receive **further academic mentoring** at their nodes and during the yearly meetings from the Internal Advisory Board, a **group of senior physicists** that includes Dr. Monica Pepe-Altarelli (also head of CERN node), Prof. Torsten Akesson from ULUND, Prof. Paula Eerola from UH, Prof. Stephanie Hansmann- Menzemer who is a deputy managing director at the UHEI, Prof. Bernard Spaan who has been Dean of the Physics Faculty of Dortmund and Prof. Giuseppe Iacobucci, the Director of the DPNC at UNIGE.

B1.3.2 Quality of the joint supervision arrangements

All academic and industrial partners play a key role in supervision of SMARTHEP, with the secondment responsibles acting as co-supervisors [17]. As described in the previous section (Sec. 1.2), the ESRs and the supervisors will create a PCDP. The supervisors will direct the research and training of the ESRs and work with them to advance RTA, and provide mentoring and support towards the ESR's future career. The supervisors will also assist the ESRs in terms of integration in their research environment, both at the local node with the help of existing students, and during secondments with the help of the dedicated secondment officer, Dr. O. Smirnova, and are the first point of contact for issues that may arise with mobility. The training officer Starovoitov [18] will act as ombudsman so that ESRs can raise issues in matter of training. During the secondments, all the ESRs will have a responsible **tutor** in the institution in which the secondment is taking place. These tutors will ensure that the internship is fruitful for the research project of the student and for their training, and will provide all the necessary resources. We have also taken care to place the secondment in the same location or country wherever possible, to ensure an effective joint supervision with maximal cross-talk, and that the overhead of ESRs moving countries (especially those with family) is reduced.

The ESR will have regular meetings with the main supervisor on a weekly basis. There will be an "open-door" policy with the members of the SC most relevant for the day-to-day work at least weekly, to review the work done and discuss next steps. This arrangement guarantees that all ESRs will receive timely supervision even when the supervisors hold positions of responsibility, and that the ESR projects are vehicles for supervisors to discuss joint research and strengthen synergies between their organizations. Whenever the students are on a secondment, the main supervisor and the SC can meet the student via videoconference. A "virtual corridor" for efficient communication between students and remote supervisors will be created in the form of instant messaging channels for each ESR project on the Mattermost platform, as described in Sec. 3.2.1. The main supervisors will report on the progress to the WP responsibles at the Executive Board meetings (see Sec. 3.2.2).

More than half of the ESRs will have a female scientist within their Supervisory Committee, which is more than the norm in the field and meets the Commission's target of 40% of the under-represented gender, and will set an example for gender balance to follow throughout the ESR's career. SMARTHEP also has dedicated diversity and inclusion officers, as discussed in Sec. 3.2.1.

B1.4 Quality of the proposed interaction between the participating organisations

B1.4.1 Contribution of all participants to the research and training program

The training program and assignment of the network events has been planned according to the research and industrial expertise of the various nodes. The training component of the network is led by UNIGE in WP2, with contributions from all beneficiaries and partners outlined in Sec. 1.2.1.

Training through research in RTA techniques is the focal point of the network. ESRs will learn to use cutting-edge tools and techniques on Machine Learning and data analysis (WP3) and hybrid computing architectures (WP4), and use them to apply RTA to decision-making (WP4) as well as to monitoring and discoveries (WP5) in research and industry, during their thesis work and secondments. Interconnections and complementarities on these topics within SMARTHEP are detailed further in Sec. 3.4. All industrial partners and most of the beneficiaries in SMARTHEP will host secondments, see List of Participants and Sec. 3.1.5.

All industrial and academic nodes contribute to the training program with specific courses during the network training events (Table 1.3), local courses and dedicated graduate schools (see Sec. 1.2.1), secondments. **Introductory ML lectures** will

[17] In the cases where the ESR is not seconded/hosted by an industry node, the ESR will have a supervisor from an industrial partner for mentoring and career advice purposes. [18] For students of the same node as the officers, the PC Doglioni will assume this role.

be taught by Sambo (KKT), Gligorov (CNRS), Pierini (CERN); **trigger and detector reconstruction** lectures will be taught by Raven (NIKHEF), Strom (UO) and Christiansen (ULUND), **HEP analysis** lectures will be taught by Albrecht (TUDO) and Voutilainen (UH), while **hybrid architecture lectures** will be taught by Lacassagne (SU), Crescioli (CNRS) and Annovi (INFN). **Advanced lectures in ML and AI** will be given by professors in ML Ustyuzhanin (Yandex responsible for CERN projects) and Louppe (ULIEGE), and Sopasakis (XIMANTIS) for use in industry. The **FPGA school** lectures and bootcamp will be taught by Crescioli (CNRS) and Boveia (CERN). The **GPU school** lectures and tutorials will be given by Martinez-Santos (USC) and Catastini (LIGHTBOX). All industrial beneficiaries and partners will contribute to the **Transferrable skills** event joint with the INSIGHTS ITN and to the **Industry and commercial applications school**. At this school, the ESRs will participate in half-day lectures given by each of the partners including: experience from transitioning from physics to industry (Dungs, from LHCb to POINT8) and vice versa (Salti, previously KKT, now UNIBO), examples of the life-cycle of a basic commercial strategy offered by LIGHTBOX, and group work in analyzing case studies (all industrial partners and beneficiaries).

All schools and events will include transferrable skills sessions, with example topics of "Getting excellent research ideas" and "Successful grant-writing" by SMARTHEP ERC grantees Pierini, Doglioni, Gligorov, Martinez-Santos and Albrecht. The researchers in SMARTHEP will also invite visiting researchers from their existing research networks, e.g. there will be lectures on sustainable and reproducible research by B. Ragan-Kelley from JupyterLab and on FAIR Data by Maria Johnsson, librarian at ULUND.

B1.4.2 Synergies between participating organisations

Participants in SMARTHEP are **experts in HEP and industrial real-time data analysis**, so the skills and goals of the network participants are complementary and enhance each other.

The SMARTHEP network includes a group of HEP researchers who have proven their ability to work and deliver large scale projects together. At the LHC, the TUDO, USC participants play a key role in the trigger and RTA of the LHCb collaboration, OSU, INFN, UO in ATLAS, CNRS, Nikhef and UHEI in ATLAS and LHCb, ULUND in ATLAS and ALICE, UH and CERN in CMS. CERN hosts all experimental collaborations and their detectors, providing an ideal ground for common training and events, as well as hands-on experience for the ESRs in physics analysis. The participation of the RTA pioneers and experts from all LHC collaborations on common analysis topics is a key feature of this network, and allows to gain the necessary momentum and critical mass to make paradigm-shifting changes in HEP. These synergies are already being exploited in obtaining joint funding and co-organizing workshops and events, such as the https://rapid2018.org and data challenge co-organized by TUDO and CNRS in 2018 on the topics of ML and RTA for detector reconstruction and industry (with Sopasakis from XIMANTIS as a keynote speaker), or the COST action CA16201 on precision HEP PARTICLEFACE with involvement of ULUND and UH as Swedish and Finnish contacts. SMARTHEP members have also played a leading role in securing funding for the “Learning to Discover” series of workshops at the Institut Pascal in Orsay. The first workshop in the series (July 15-26 2019) will explicitly focus on RTA, in particular how best to exploit hybrid architectures to enhance RTA capabilities in research and industry; bringing together SMARTHEP consortium members, other LHC collaborators and industry partners. SMARTHEP is the perfect opportunity to enhance and extend those collaborations to a wider network with a long-term perspective, bringing advancements beyond the state of the art for RTA at the LHC.

These advancements require specific expertise from computer science and data processing that goes beyond HEP. The non-academic partners have been chosen so that the ESRs can learn hands-on from RTA and and decision making in the topics of transport, finance and IoT, while cross-pollinating industry applications with the skills acquired from the analysis of LHC data. Experience in applications of machine learning for RTA is provided by ULIEGE, UNIBO, IBM, XIMANTIS and POINT8. In particular, IBM can contribute to solving one of the key challenges for RTA in both HEP and fraud detection: the need to have understandable and reproducible decision-making processes in real time, using rule induction techniques. Real-time sensors and hybrid architectures are a specialty of LIGHTBOX and KKT, in synergy with the expertise of UNIBO and SU in terms of computing on resource-constrained platforms and edge computing. In all those topics, non-academic beneficiaries and partners offer ESRs and HEP researchers the perfect opportunities to exploit ideas and algorithms created for fundamental research in a commercial context, promoting the ESR advancement in both fields and solving challenges that are crucial for both academia and industry.

The non-academic beneficiaries also have more specific connections beyond academia as depicted in Fig. 2 and in the ESR descriptions, so that the industrial partners can benefit from each other’s expertise and advance their business. Those connections will be discussed among partners and beneficiaries in topical

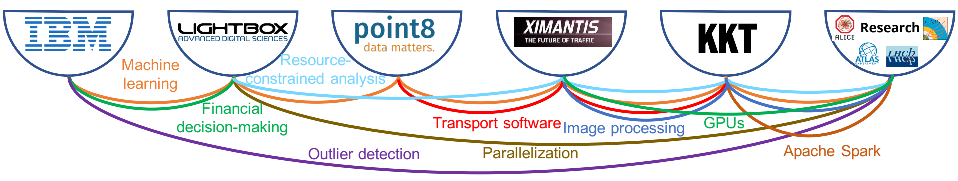

Figure 2: Synergies between industry and academic participants.

meetings during the yearly network events.

B1.4.3 Exposure of recruited researchers to different (research) environments, and the complementarity thereof

SMARTHEP will expose all ESRs to environments and challenges of HEP and industry. This will create a generation of interdisciplinary problem-solving researchers. Within their project, ESRs with a computing/hardware background will work on LHC data analysis, and ESRs with a HEP background will work on the techniques and methodologies of data science, software and hardware programming. All ESRs will be exposed to commercial applications of RTA, by conducting their PhD research in an industry beneficiary, or through secondments to industry, as well as via mentoring and network-wide events.

B2. IMPACT

B2.1 Enhancing the career perspectives and employability of researchers, contribution to their skills development

ESRs will gain a skillset that spans from research to industry, in an European context, as the proposed network builds on the foundation of **interdisciplinary, public-private** exchanges and **international scientific collaborations** To describe the impact of the research and training on the ESR careers, we follow the headers of the Seven Principles of Innovative Doctoral Training[19].

Research excellence, attractive institutional environment. The research and peer-reviewed publications on topics that are crucial for the advancement of HEP, stemming from the RTA centered research program of SMARTHEP, as well as a Ph.D. degree from renowned European universities, will provide the ESRs with the best possible ground towards a career in academia. All institutes in SMARTHEP have long experience in supervising and bringing students to graduation of master and Ph.D. degrees, and employ world-renowned scientists that are leaders in RTA and can serve as supervisors as well as role models.

Interdisciplinary Research Options, Exposure to Industry and other relevant employment sectors. Training in large-scale Data analysis from HEP experiments, especially those trained in ML and data science, already provides a good ground for employability[20]. Bridging the gap between academia and industry therefore remains essential for most physics graduates[21]. ESRs will not only be exposed to their own topic of research at their local institute, but they will also be seconded at other institutes and industrial partners with practical applications that are sufficiently diverse but still match their research topic, leading to a consistent thesis project as well as a broader expertise for the ESRs at the end of their studies.

International networking. The complexity and cost of HEP research has pushed scientists to form large, international collaborations. Scientists have an ideal ground to do so at CERN: the idea of trans-national scientific collaboration promoting a unified Europe is at its heart. All SMARTHEP ESRs research focused on developing RTA within the CERN experiments, and CERN will host PhD students and secondments. By virtue of this, the ESRs will be immersed in an international environment that trains them to work in teams and present their work regularly. It also encourages the ESRs to consider Europe as a single, open labour market, so that they can either continue their research quest towards a deeper understanding or nature or convert their ideas into products and services for a thriving knowledge-based European economy and society.

Transferable skills training. The academic and industry training foreseen for the ESRs in one of the most data intense fields prepares them for careers in academia, industry, or both, following the example of their supervisors. The training program of SMARTHEP is crucial for this purpose, and includes a significant portion of transferrable skills training described in Sec. 1.2.

Quality Assurance. Both ESRs, and the SMARTHEP research program, will be continuously reviewed through feedback of their local supervisors and peers. Further internal quality assurance is provided by the other members of the network, before coming to the peer-review and selection of papers and conference talks, as well as the exploitation of the research of products developed together with the industrial partners. This combination will allow SMARTHEP ESRs to complete an excellent PhD degree that will be recognized by future employers. The secondments to the industrial partners will allow the ERSs to gain valuable, documented work experience that will boost their employability and recruitment possibilities, both within and outside the the companies of the network and prepare them for joining the European work force towards successful careers.

B2.2 Contribution to structuring doctoral/early-stage research training at the European level and to strengthening European innovation capacity

B2.2.1 Contribution to structuring doctoral training at the European level

The academic and industrial training offered by SMARTHEP greatly benefits ESRs by widening their career choices beyond the academic model of student-postdoc-professor. Both SMARTHEP recruitment and **training** proceeds in a diverse, inclusive, and supportive environment with the help of mentors and specific officers, complementing taught courses at the beneficiary institutes with academic research and non-academic industrial secondments. Such a diverse training program mandates an attentive SC and the use of a credit system linked to the ECTS system, to quantify and evaluate ESR training progress (Sec. 1.2). Both serve as example for the structuring of early-stage research training within SMARTHEP.

[19] ERA Steering Group Human Resources and Mobility, European Doctoral Training Principles, 2011 [20] Wired Jan 2017: Move Over, Coders? Physicists Will Soon Rule Silicon Valley [21] P. Heron, Preparing physics students for 21st-century careers, Physics Today 70, 11, 38 (2017);

SMARTHEP training happens within and outside the consortium, blending with existing training programs hosted by beneficiaries and partners, as well as with ongoing ITN networks (e.g. organizing joint training with INSIGHTS). As well as providing complementary training for ESRs, these collaborations in training offer an excellent platform to share best practices and seek complementary funding sources beyond this ITN, increasing the internationalization of the participating organizations. Discussions have already started among the network participants on how to share training and find funding beyond this ITN to attract and retain more talented students from abroad. The participation of the network to ISOTDAQ and MLHEP, the most important school of HEP data acquisition and ML, goes in this direction, by disseminating network results and techniques and at the same time allowing ESRs to learn the state of the art.

SMARTHEP links training in HEP with commercial applications within the common ground of RTA, so that the non-academic sector strongly contributes to the increased societal and economic relevance of the outcome of the ESR research and training path. The majority of ESRs will spend a minimum of three months seconded to an industrial beneficiary or partner, and will explicitly use their academic training to further the development of industrial products; in turn, the industrial training they receive will boost their academic research. Within SMARTHEP, the non-academic sector is placed at the heart of the training program in a concrete and practical way, so that ESRs employ techniques developed for research in industry and vice versa. This is not yet the norm in physics graduate courses, even though studies show a need for this kind of interdisciplinary and intersectoral mobility experience[22]. Moreover, history shows that such interdisciplinary cross-pollination of methods and problem-solving skills strengthens innovation (e.g. Bell Labs in the USA). We hope the success of SMARTHEP, if funded, and of H2020 ITN programs across Europe will encourage embedding non-academic training in PhD degrees.

B2.2.2 Contribution to strenghtening the European innovation capacity

The research and industry goals in SMARTHEP be met using innovative techniques that are the focal point of a European-wide ESR education, benefitting both the European higher education system and Europe's innovation and growth potential.

The economic importance of physics in European industries overall has been detailed by the CEBR[23] in a report commissioned by the European Physical Society. Industries for which the use of physics in terms of technologies and expertise is critical to their existence generated a turnover of € 3.8 trillion in 2010, 15% of the total turnover Europe's economy. Moreover, the productivity per employee in the physics-based industries largely outperforms that of other industries.

As noted by a recent report by the EU's High Level Expert Group on the European Open Science Cloud[24], expertise in Data Science is a skill needed to strenghten innovation in Europe, and one that SMARTHEP ESRs will gain. SMARTHEP will prepare a well-trained addition to the European workforce that will give a competitive edge to the academic institutes and industries where the ESRs will undertake their training, as well as those where they will be employed after their PhD. This will be increased by the strength of the network as a whole, acting coherently to improve the deployment of RTA techniques. In the data-rich environment that will dominate both HEP and industry, making decisions fast and efficiently becomes a priority to make an impact and be at the forefront of innovation and developments.

The LHC is at the forefront of big data in the scientific domain: Computing sites that are part of the Worldwide LHC Computing Grid need around 1 exabyte to store data from the LHC experiments after the trigger selection[25]. While LHC data may not be as big as for large companies, cost-effectiveness is crucial for research due to the more limited budget allocated. The codebase of the LHC experiments, more than 50 million lines long[26] is larger than all but the biggest commercial projects. For these reasons, the LHC research ecosystem provides an ideal opportunity to develop new methods of analyzing big data in real-time, and to train ESRs in topics matching the emerging trend of **real-time big data analytics**[27]

B2.3 Quality of the proposed measures to exploit and disseminate the project results

B2.3.1 Dissemination of the research results

In order for SMARTHEP results to make an effective impact within and beyond the consortium, and to ensure a high visibility for ESRs, we identify three distinct audiences for the dissemination strategy. The first is composed of the **scientists members of the LHC collaborations** that the ESRs are connected to. In this context, intermediate work can be shown by the ESRs, normally after discussion and preview by the supervisor. In addition, the ESRs will present their work regularly within SMARTHEP at the occasion of the yearly meetings, as all four LHC collaborations allow the presentations of intermediate results within the context of regional or training network meetings after a brief review process.

When the results are mature, they will be made public to **the wider HEP and Data Science communities**. This step will happen after internal peer-review from the collaborations for physics results, and after the necessary exploitation measures have been taken in case of algorithms with a commercial value. SMARTHEP results will be published in high-impact Open Access

[22] Roach M, Sauermann H (2017) The declining interest in an academic career. PLOS ONE 12(9): e0184130, `http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0184130` [23] The importance of physics to the economies of Europe, Centre for Economics and Business Research Ltd, London, 2013 [24] "500,000 data scientists needed in European open research data", A. Offermann, `https://joinup.ec.europa.eu/news/500000-data-scientists-need` [25] The HEP Software Foundation, A Roadmap for HEP Software and Computing R&D for the 2020s, ArXiv Document Server [26] Information Is Beautiful, Codebase infographics, accessed online 01/2019 [27] A comprehensive guide to real-time big data analytics.

journals of the most closely corresponding research field, whether physics or Data Science. The physics publications produced within the ALICE, ATLAS, CMS and LHCb collaborations fall under the CERN regulations on intellectual property. CERN publishes following the Open Access conditions, as defined by the SCOAP3[28] initiative, adopting the Creative Commons Attribution. The CERN publication policy aims to exercise the copyrights to permit the widest possible dissemination and use of the obtained results, and all SMARTHEP publications will follow the same publication philosophy. Examples in which SMARTHEP members have regularly published are PRL[29], PLB, JHEP, JMLR, ML, IEEE PAMI. All publications will also be made available on the CDS and arXiv preprint servers. Each ESR will be responsible for the publication of 1-2 journal papers and 1-2 conference reports during their PhD. SMARTHEP members have extensive experience preparing publications in the most qualified peer-reviewed journals of their respective fields and will consequently supervise and train ESRs to disseminate their results. One of the priorities of SMARTHEP is to train ESRs in reproducible and sustainable research. Software toolkits will be published on Zenodo, so that they can be cited. Wherever appropriate, Jupyter notebooks and Docker images to run data analysis will be published together with research papers. SMARTHEP results will also be presented in international conferences, e.g. EPS, ICHEP, Moriond EW&QCD, NIPS, ICML, AAAI, ECML, ICLR, AISTATS, CHEP, CDSR, IEEE Real Time Conference, etc.

Finally, reports and whitepapers delivered within the research WPs, will summarize the overall WP results with respect to the initial Network objectives, and the recommendations made towards adopting RTA and the computing resources needed to do so. The target audiences here are the **EU Commission, the European researchers subscribed to the EURAXESS community**, who may find interesting links and connections between their research and the research undertaken in SMARTHEP and want to collaborate, as well as policymakers and research councils who can be pointed to whitepapers and recommendations, in the same fashion as for the HEP Software Foundation Whitepaper[30] that encouraged DOE to fund the IRIS-HEP institute for computing in HEP in the US. Members of SMARTHEP (Sokoloff, Gligorov, Boveia, Albrecht and Doglioni) are active members of the HSF and hold positions of responsibility. Boveia and Doglioni also have experience being the main organizers and editors of the Dark Matter Forum and Dark Matter Working Group whitepapers, with recommendations on Dark Matter benchmark models for the entire LHC theory and experimental community. Together, SMARTHEP is equipped with key people to achieve this dissemination goal.

The responsible for the dedicated Dissemination and Outreach Work Package (Petersen, CERN) will assist the network members (ESRs and supervisors) on publication matters. He will also ensure that all results of SMARTHEP and their online links are highlighted on the SMARTHEP web page, social media and local institute press releases.

B2.3.2 Exploitation of results and intellectual property

The main academic results of SMARTHEP are software, peer-reviewed papers and whitepapers.

As mentioned in the previous section, the physics publications produced within the ALICE, ATLAS, CMS and LHCb collaborations fall under the CERN regulations on intellectual property. Similar terms apply to the data and software tools generated within the academic institutions and the CERN laboratory. SMARTHEP will conform to the open access policies of the four LHC collaborations when it comes to the data collected by these detectors.

The developed software will be released as described in Sec. 3.2.7, and used in the ALICE, ATLAS, CMS and LHCb collaborations and disseminated for further use in research. It be licensed under Open Source licenses such as Apache/MIT/BSD/GPL depending on the use case, ensuring free usage and distribution. An exception applies to results produced by ESRs working in the private companies. In this case the internal rules of the companies concerning the Intellectual Property of the work need to be respected to guarantee its potential commercial exploitation. Preliminary arrangements with the concerned companies have been made, so that the obtained development are published in journal papers, while their actual implementation remains the Intellectual Property of the company. More details of IPR can be found in Sec. 3.2.7.

The main commercial results of SMARTHEP are the improved XIMANTIS traffic prediction application, fleet software for edge computing and mobile platforms, software for real-time anomaly detection through automated inference and rule induction in IBM, and Internet-of-Things sensors for industrial production processes in Lightbox. Both beneficiary institutes and secondment hosts will benefit from the exploitations in joint intellectual property agreements. As an example, the adoption of predictive maintenance in industrial production processes thanks to the Internet-of-Things sensors development during the Lightbox secondment, is estimated to on average reduce maintenance costs by more than 25%, reduce breakdowns by more than 70%, reduce downtime by more than 35% and increase productivity by more than 20%. While we cannot quantify the commercial impact of all industrial products in the same way, we are convinced that the other secondment projects which the ESRs will perform with their industrial partners, and the innovation of methods which SMARTHEP will more generally foster, will lead to significant additional economic value.

---

[28] SCOAP3, SCOAP3 Sponsoring Consortium for Open Access Publishing in Particle Physics. [29] We are aware that PRL is not Open Access or SCOAP3 compliant, but CERN articles are treated as a special case by PRL and **are** Open Access, and all SMARTHEP publications sent to PRL will be in this category. [30] http://hepsoftwarefoundation.org/activities/cwp.html

We have preliminarily reserved 15000 EUR from the management budget [31] for filing a possible patent on one of these commercial deliverables, to boost the exploitation of the most promising algorithms and toolkits in SMARTHEP. The use of this fund for patent filing will have to be approved unanimously by the Supervisory Board.

B2.4 Quality of the proposed measures to communicate the project activities to different target audiences

B2.4.1 Communication and public engagement strategy

Reaching audiences with a direct interest in SMARTHEP research is covered in Section 2.3.1. It is equally important to inspire the general public, young generations of students and their educators, the entrepreneur and industry community as well as stakeholders and politicians with SMARTHEP's basic research and novel techniques whose commercial applications impact everyday life.

Communication: The ALICE, ATLAS, CMS and LHCb experiments, CERN, and academic institutions involved in SMARTHEP have **public web pages** where the experiments main results are presented to large audiences. SMARTHEP already has a webpage (`www.smarthep.org`), describing its members and their activities. If the network is successful, the page will be expanded with the ongoing activities, goals, achieved results and latest news with dedicated general public target. It will also continue being updated as the members engage in further collaborations beyond this ITN. The webpage will contain a link to a blog curated by ESRs who would like to discuss their research activity in the style of http://www.quantumdiaries.org, to allow the general public to have an insight into the life and work of young researchers. Here the ESRs can benefit from the experience of the INSIGHTS ESR seconded at ULUND who is already curating such a blog. **Social media** are an important part of SMARTHEP communication strategy: many of the network researchers are already active on Twitter (see e.g. @CatDogLund for the PC's account) and the network will have its own Twitter handle. We plan to run a scheme where an ESR curates the account for a period of time in the same fashion as @CMSVoices or @Sweden, with tweets approved by the PC and the WP7 responsible. Highlights and milestones of SMARTHEP will be posted from this account. SMARTHEP researchers will provide regular information to their institutions's press offices, and be proactive in seeking contact with **local newspapers and TV channels**, through press releases and articles in accessible language (e.g. we foresee an upcoming article on RTA in The Conversation website[32]). The **logo** designed for SMARTHEP and the H2020 logo will provide visibility to the network by being included on presentations in conferences. Stakeholders and politicians will be reached through work with the Embassies of the network partners, in a similar fashion as what done in the RAPID2018 workshop supported by both the French and German embassy and was attended by the embassy responsibles. Examples of platforms for this activity are strategic partnerships between countries member of SMARTHEP that are related to the network's deliverable, e.g. between France and Sweden on innovative and green solutions in the transport sector.

Outreach: Outreach activities are an integral part of SMARTHEP training.The main innovative outreach element in SMARTHEP is a **RTA data challenge** organized by Ustyuzhanin (CERN). This is a citizen science project[33], where open problems about RTA designed by the ESRs and their supervisors are published on an online platform, and prizes are offered to those solving them in the most efficient way. CERN and other SMARTHEP participants have experience in citizen science and data challenges with e.g. the HiggsHunters project (Oxford U. and ULUND) and the RTA-tailored RAPID2018 workshop (TUDO and CNRS). Since HEP is mainly publicly funded, SMARTHEP members are already involved in outreach and educational activities such as **talks in high schools, conferences on popular science, summer schools, mentoring** and similar programs. They are also active in the EU-organized and funded events such as the **European Researchers Night**. These activities will be used as a platform for giving visibility to SMARTHEP's multidisciplinary activities. **Open lectures** to, as well as Q&A sessions with, the general public and local schools will be included as a part of SMARTHEP-wide events. At the yearly meetings, we will organize **"Science on tap"** events following the successful example of the University of Hamburg and DESY (with which ULUND has a strategic collaboration)[34], where the ESRs present will organize short talks in local venues in an informal setting, and get to know each other and the attendants. In addition, SMARTHEP members have extensive experience with the **International Masterclass** program and **World Wide Data Day**, which exposes tens of thousands of high-school students worldwide to HEP concepts and techniques. As part of WP7, we will create two new RTA-based Masterclass and WWDD exercises that can be chosen by any institute, moderated by SMARTHEP members. In the same vein, it is vital to convey to undergraduate students in multiple disciplines (physics, mathematics/statistics and economy) the relevance of the network activities and the opportunities that the job market has for people that have completed a Ph.D. in the multidisciplinary context created by SMARTHEP. For this reason, ESRs will be encouraged to help with CERN visits from local schools when seconded or on field trips there. All ESRs will be expected to include at least one regular local outreach activity (e.g. chool and university visits; mentoring of younger students with a special eye on minority communities) within their PCDPs. All public engagement activities undertaken by ESRs will count as credits towards their training. To

[31] Quote as per discussion with LU Innovation, for filing first a national and then an international patent [32] The Conversation is an independent research-oriented news source, `https://theconversation.com/`. [33] Citizen science engages the public in the creation of scientific results, reducing the gap between the scientific community and the general public. [34] Science on tap 2017 edition, `http://www.cui.uni-hamburg.de/en/events/science-on-tap-2017/`

reach an audience with an interest beyond academia and inspire further collaboration, we foresee that the ESRs and the industrial supervisors will participate to local innovation events [35] and career fairs, presenting the results of SMARTHEP on RTA applications and the model of academic and non-academic training for PhD students. This will also be beneficial for the exposure and recruiting strategies of the industrial partners.

B3. QUALITY AND EFFICIENCY OF THE IMPLEMENTATION

B3.1 Coherence and effectiveness of the work plan

B3.1.1 Work Packages description

SMARTHEP research topics and corresponding WPs have been introduced in Sec. 1.1.2. The ESRs will work on projects which fall within different research WPs, strengthening SMARTHEP into a coherent whole. RTA is the conceptual and technical challenge which permeates all research work packages and motivates the network.

WP1	Management	**Lead Beneficiary**: ULUND	**Duration**: 1-48	ESR: 1 elected
Objectives: Create and maintain a high level of collaboration between the SMARTHEP consortium members to ensure the optimal functioning of the network, and report to the EU on network activities and progress.				
Description of Work and Role of Partners: Doglioni as Project Coordinator (PC) will oversee the smooth running of the planned research and training activities for the network. Lund University has been chosen to coordinate SMARTHEP as a large institution with dedicated support and experience with projects of this kind, and due to Doglioni's experience in coordinating HEP communities and groups within and beyond ATLAS (e.g. LHC Dark Matter Working Group, more than 300 participants and HEP Software Foundation trigger and reconstruction WG). Doglioni is an alumna of the Collegio Universitario Lamaro Pozzani, funded by the Italian National Federation of Holders of the Order of Merit for Labour and hosting up to 15 selected Italian students per year, complementing their regular university education with courses and lectures in entrepreneurship, law, IT and languages. This gives her an unique perspective on cross-talk between academia and industry that started developing as early as her undergraduate education. Administrative support to the PC will be hired, at the level of at least 50% FTE. The PC will keep close contact with other WP coordinators to monitor the execution of tasks, and compliance to the defined milestones. The PC's first task will be to convene discussions leading to the signature of the Consortium Agreement (CA). Management will coordinate the hiring of all ESRs together with a dedicated recruitment officer.The PC is responsible for preparing and chairing the consortium administrative meetings in the EB and SB, for the preparation of reports for the EU together with the project manager, and for overseeing the election of ESR representatives to the EB and SB. The PC and the relevant WP coordinators oversee the organization of the planned workshops and conferences The PC is responsible for facilitating communication between the nodes, building lasting collaborations between the nodes, enabling knowledge transfer and forming a legacy of future research collaborations.				
Deliverables: **1.1** (month 2) Hiring of a dedicated project manager (month 1); **1.2** Signature of the CA by all parties; **1.3** (month 3) Launch of the website and social media accounts as outward-facing communication and organization tools; **1.4** (month 3, 8) Advertisement of the recruitment and recruitment completion, including ESR declarations; **1.5** (month 3,9,23,36,48) Prepare brief reports for each SB meeting, incl. mandatory annual and mid-term progress report to EU ; **1.6** (month 3,9,23,36,48) Joint with WP2, collect documentation, reports and feedback to be discussed at the upcoming SB, following all Network events and schools.				

WP2	Training	**Lead Beneficiary**: UniGe	**Duration: 1-48**	ESR: All
Objectives: Organisation of the Network-wide training events, overview of the personal training for each ESR, see Sec. 1.2.				
Description of Work and Role of Partners: The WP2 coordinator is Sfyrla from UniGe, with experience in student supervision, public lectures and course organization. The WP2 coordinator oversees the preparation of the Personal Career Development Plans and follows up with the supervisors for the intermediate and final reports Prof. Sfyrla oversees the organization of the lectures for the development of technical, research and transferable skills (see Sec. 1.4.1) taking place at the Network-wide events. The WP2 coordinator makes sure the lecturers give training of excellent quality and that adequate documentation is provided to the ESRs, and provides advertisement and follow-up material together with the PC. Short lecture proceedings will be prepared in advance of the lectures by the speakers so that they can be disseminated without delay following the events (as for example with the CHEP conference series). All nodes and partners will benefit and contribute to the training of the ESRs, by playing an active role through supervision of PhD programs, research and industry projects that will train excellent scientists with a wide range of skills and experiences. Together with the responsibles for WP7, Sfyrla will ensure the quality and feedback of the presentation of ESRs to SMARTHEP events.				
Deliverables: **2.1** (month 9,23,36,42) PCDPs for each ESR, intermediate and final monitoring; **2.2** (month 10,15,27-29,38) Design and organization of network-wide schools together with responsible beneficiaries; **2.3** (month 9,23,36,42) Presentations from all ESR to the SMARTHEP events and at the final conference;				
2.4 (month 44) ESRs with three-year PhD duration receive a degree.				

WP3	ML & advanced data analysis	**Lead Beneficiary**: CNRS	**Duration: 8-48**	ESR: All
Objectives: Deployment of advanced Machine Learning (ML) and data analysis techniques to enable real-time analysis.				
Description of Work and Role of Partners: The WP3 coordinator is Gligorov from CNRS, who led the first large-scale implementation of real-time ML at an LHC experiment in 2011, led LHCb's HLT during 2014 and 2015, oversaw LHCb's physics programme as deputy physics coordinator during 2016 and 2017 and is currently leader of the Real Time Analysis Project in LHCb (consisting of 30 institutes and around 50 FTE). The WP3 coordinator is responsible for the coherence and interaction of the ESRs investigating ML techniques in RTA throughout SMARTHEP, for their development and their application to HEP and commercial use cases according to the network-wide strategy described in Fig. ??. He will also ensure that code is written alongside documentation that makes the software useful and usable, and that frameworks and toolkits are made public in a timely manner and respecting IP clauses in case of commercial exploitation. WP3 will produce novel algorithms to reconstruct objects and events in real-time (with the know-how of NIKHEF and CERN), and develop ML techniques for event reconstruction, fast data analysis and outlier detection (exploiting the expertise of ULIEGE, IBM, KKT and UNIGE). The algorithms for both reconstruction and ML will also be benchmarked and optimized in the projects hosted at NIKHEF and CERN.				
Deliverables: *For this and other research WP deliverables, we follow the implementation plan in 1.1.1, where the ESRs first gain an overview of the state of the art, then developing new techniques, and finally disseminating them as legacy of SMARTHEP.* **2.2** (joint with WP1, WP2) design, organization and documentation of the contributions to the network ML and MLHEP schools **3.1** (month 13) Whitepaper on the state of the art on ML for real-time analysis, detailing implementation and deployment, capitalizing on the attendance of the MLHEP school in month 24; **3.2** (month 31) Collection of ML algorithms and software toolkits to be exploited for the research objectives in WP5 and WP6; **3.3** (month 34) . Review paper collecting description and documentation of techniques for ML in RTA.				

[35] e.g. Ideon Science Park Innovation Challenge, `https://ideon.se/event/lund-innovation-challenge/`

WP4	Hybrid architectures	Lead Beneficiary: SU	Duration: 8-48	ESR: 4-9, 14

Objectives: Study and adoption of hybrid computing architectures to enable RTA.

Description of Work and Role of Partners: The WP4 coordinator is Lacassagne from SU. He is a system architect with extensive experience in the benchmarking and use of hybrid architectures. The WP4 coordinator is responsible for the coherence and interaction of the ESRs developing and testing code for hybrid architectures. He is responsible for the publication and exploitation of successful deliverables related to hybrid architectures, as well as many studies that demonstrated that standard architectures could be improved for the purpose of RTA (see e.g. Lemaitre, Lacassagne, Batched Cholesky Factorization for tiny matrices, DASIP 2016). The work is divided in three tasks corresponding to the research objectives: the use of FPGAs (e.g. track triggering for ATLAS, expertise of OSU, UO, INFN), the use of GPUs for speeding up parallel algorithms in industry and HEP (USC), and employing parallel and multithreaded algorithms (LIGHTBOX). Together with the partners, the WP4 coordinator also oversees the training program on each specific architecture, and ensures the quality of lectures and their proceedings with the coordinator of WP2. He also makes sure that code written for non-standard architectures satisfies high documentation standards.

Deliverables: **2.2** (joint with WP1, WP2) design, organization and documentation of the FPGA and GPU schools.
4.1 (month 20) Whitepaper on the state of the art on hybrid architectures in real-time analysis, capitalizing on attendance of network FPGA/GPU schools;
4.4 (month 38) Software toolkits and hardware improvements in HEP (e.g. FTK for ATLAS, GPU for LHCb reconstruction);
4.2 (month 25) Lightbox software to optimize parallelization of financial transactions and associated publications;
4.4 (month 40) Review paper collecting advancements in optimization of hybrid architectures for the LHC trigger systems.

WP5	Real-time decision making	Lead Beneficiary: TUDO	Duration: 8-48	ESR: 5, 7, 8, 10-14

Objectives: Applying real-time analysis to decision making in physics and society.

Description of Work and Role of Partners: WP5's goal is to enable fast and efficient decision making with RTA, in physics through the use of the trigger systems, and in society to improve safety and efficiency of transport in ways that would not be possible without RTA. WP5 is coordinated by Albrecht (TUDO as main beneficiary node) with Sopasakis as co-coordinator (Ximantis). They have been chosen to fill this role for their complementary expertise in decision-making in HEP and industry that are crucial for SMARTHEP. They will ensure that the RTA techniques enabled by WP3 and WP4 are applied to advance both HEP and industry by the various ESRs working on the experiment trigger systems, that the exchange between academia and industry is fruitful in both directions through cross-pollination of techniques, and that the results are documented in peer-reviewed papers.

Deliverables: **2.2** (joint with WP1, WP2) organization of trigger contributions to network events and ISOTDAQ schools.
5.1 (month 14) Review of the state of the art of the triggers of LHC collaborations, compiled by ESRs will prior to their physics analyses capitalizing on attendance of ISOTDAQ school and network events;
5.2 (month 36) Improved XIMANTIS app using novel ML techniques (e.g. hybrid networks) and associated publications;
5.3 (month 31) Software for LHC trigger upgrades for Run-3 data taking;
5.4 (month 29) Client software for optimization of transport and logistics with POINT8 and associated publications;
5.5 (month 42) Review paper collecting physics results using trigger selection improvements, including summary of publications on dark sectors, LFV/LFU and precision measurements.

WP6	Real-time monitoring and discoveries	Lead Beneficiary: IBM	Duration: 8-48	ESR: 1-4, 6, 7, 14, 15

Objectives: Applying real-time analysis to monitor complex systems and discover anomalies, in physics and society.

Description of Work and Role of Partners: The goal of WP6 is to employ real-time analysis to detect novelty or anomalies while monitoring complex systems and streams of data. These data streams range from LHC collision events (ULUND), to financial transactions (IBM), to data from vehicle dashboard cameras (KKT), to sensor data from industrial processes (LIGHTBOX). A sub-goal that is novel to SMARTHEP and essential to introduce such techniques in HEP trigger systems is the accountability and reproducibility of the algorithms employed, developed in ESR3. For this reason, IBM is chosen as the lead beneficiary of WP6 with De Sainte Marie as main coordinator, given his extensive experience in supervision of student projects and his expertise on symbolic knowledge systems. WP6's academic co-coordinator is Pierini (CERN), one of the pioneers of anomaly detection in LHC experiments. WP6 will work in close collaboration with WP3 to design new algorithms and combine the best of both symbolic knowledge and numerical algorithms towards application to HEP triggers and society. The deliverables of WP6 match the research objectives and include applications both in HEP and in the commercial sector, as the algorithms developed can be ported to the different kinds of data. Aided by the WP7 coordinator, by LU Innovation from the PC side and by the H2020 IPR helpdesk, WP6 coordinators ensure that these commercial deliverables are disseminated and documented after their exploitation, and correctly handled in terms of IP and EU GDPR.

Deliverables: **2.2** (joint with WP1, WP2) organization of the non-academic training and Industry and career development schools.
6.1 Review of the state of the art of fully RTA searches in HEP and recommendations for improvements (companion of **5.1**) (month 17) ;
6.2 (month 31) Software enabling advanced fully RTA-based searches at LHC;
6.3 (month 42) Algorithms for fraud detection and HEP triggers in IBM and associated publications;
6.4 (month 42) Toolkits using ML and AI for real-time in-fleet monitoring within KKT and associated publications;
6.2 (month 31) Software for sensors for Internet-of-things and industrial process optimization in LIGHTBOX
6.3 (month 42) Review paper collecting physics results using fully RTA-based analyses, including summary of publications on dark matter mediators, Higgs boson, heavy ion physics.

WP7	Outreach and dissemination	Lead Beneficiary: CERN	Duration: 1-48	ESR: All ESRs.

Objectives: Relay SMARTHEP scientific activities to the general public, monitor delivery of results as journal papers, ensure ESR visibility in conferences.

Description of Work and Role of Partners: WP7 is detailed further in Sec. 2.4.1. CERN is chosen as the lead beneficiary of WP7 given its extensive experience in communicating with the public, with Petersen as responsible. The WP7 coordinator organise the communication of SMARTHEP to both the general public and the scientific community and ensure that all ESRs and supervisors take part in the effort. Together with the PC, the WP7 coordinator runs the dedicated communication portal `www.smarthep.org`, supported by ESR blogging and social media activities. Dissemination to the scientific communities is achieved by poster and talk contributions to conferences. The WP7 coordinator will delegate members of the network to support the ESR supervisors and monitor the quality of the material presented by the ESRs and ensure they benefit from an outstanding international visibility. Since the SMARTHEP data challenge is a central piece of our outreach strategy, Ustyuzhanin from CERN and Yandex will be responsible for part of this effort benefitting from his expertise and work. Further outreach to the general public will take the form of visits to schools, guided visits to CERN facilities, public lectures and "science on tap" at Network events, while stakeholders and politicians will be reached by work with the dedicated research officers in the nodes national embassies. These will be complemented by a specific SMARTHEP IPPOG Masterclass exercise. These tasks will ensure the general public grasps the impact of SMARTHEP on academia and HEP, industry and everyday life and actively participates in it through the data challenge. All academic and industrial partners will take an active part in the WP7 activities, and will simultaneously host online the International Masterclass exercises and World Wide Data Day (WWDD) activities, including on the International Day of Women in Science day.

Deliverables: **1.3** (joint with WP1) Launch of `www.smarthep.org` and social media as platform for dissemination, communication and outreach;
7.1 (month 14) Data challenge website online and open to the public;
7.2 (month 23, 36, 48) Reports on outreach to the general public and on SMARTHEP Masterclass/WWDD exercise.
7.3 (month 24, 36, 48) Report on presentation of results at international conferences;
7.4 (month 24, 36, 48) Report on publication of results in peer-reviewed journals;

Table 3.1a Work package description for each work package.

B3.1.2 List of major deliverables

The following tables list the deliverables of the project. Deliverables are assigned to ESRs in the next section taking care of the feasibility of the work plan (e.g. that they are only editing one of the whitepapers of their WP that are delivered on a similar timescale).

Scientific Deliverables

Deliverable Number & Title		WP	Lead Benef. Short Name	Type & Dissemin. Level		Due Date
3.1	Whitepaper summarizing the state of the art in ML techniques for RTA in SMARTHEP use cases and best practices	3	CNRS	R	PU	13
5.1	Review of the state of the art of the triggers of LHC collaborations and best practices	5	TUDO	R	PU	14
6.1	Whitepapers on the state of the art of fully RTA-based searches and anomaly detection in HEP and best practices	6	IBM	R	PU	17
4.2	LIGHTBOX software to optimize parallelization of financial transactions, associated publications	4	CNRS	OTHER	CO	25
5.4	Client software for optimization of transport with POINT8 and associated publications	5	TUDO	OTHER	CO	29
4.1	Whitepaper on the state of the art on hybrid architectures in RTA in SMARTHEP use cases	4	SU	R	PU	20
6.2	Software for sensors for Internet-of-things and industrial process optimization in LIGHTBOX	6	IBM	OTHER	CO	31
6.5	Update of KKT app with smart insights for drivers	6	IBM	OTHER	CO	31
3.2	ML algorithms and software enabling RTA and its benchmarking	3	CNRS	OTHER	CO	31
5.3	Software for LHC trigger upgrades for Run-3 data taking	5	TUDO	OTHER	CO	31
6.2	Software enabling advanced fully RTA-based searches at LHC	6	IBM	OTHER	CO	31
3.3	Review paper collecting ML and data analysis advancements from ESR projects	3	CNRS	PDE	PU	34
5.2	Improved XIMANTIS app (e.g. using hybrid/Attention ML networks) and associated publications	5	TUDO	OTHER	CO	36
4.4	Software toolkits and hardware improvements in HEP (e.g. FTK for ATLAS, GPU for LHCb)	4	CNRS	OTHER	CO	38
6.4	Toolkits for real-time fleet monitoring within KKT and associated publications	6	IBM	OTHER	CO	42
6.3	Algorithms for outlier detection for fraud detection and HEP discoveries in IBM and associated publications	6	IBM	OTHER	CO	42
4.4	Review paper collecting advancements in optimization of hybrid architectures for LHC triggers and related industry use cases	4	CNRS	PDE	PU	40
6.3	Review paper collecting ESR physics results using fully RTA-based analyses	6	IBM	PDE	PU	42
5.5	Review paper collecting ESR physics results using trigger selection improvements	5	TUDO	PDE	PU	42

Management, Training, Recruitment and Dissemination Deliverables

Deliverable Number & Title		WP	Lead Benef. Short Name	Type & Dissemin. Level		Due Date
1.1	Project manager hired at the PC's institute	1	ULUND	ADM	CO	2
1.2	Signature of Consortium Agreement by all parties	1	ULUND	ADM	CO	2
1.3	Launch of the website and social media accounts	1,7	ULUND	PDE	PU	3
1.4	Advertisement, selection and recruitment of all ESRs	1	ULUND	ADM	CO	3, 8
1.5	EB/SB meeting minutes and EU reports	1	ULUND	R	CO	3,9,23,36,42,48
2.1	Preparation, network approval and monitoring of PCDP	2	UNIGE	OTHER	CO	9,23,36,42
2.2	Design and organization of network-wide schools together with other responsibles	2-6	UNIGE	OTHER	CO	10, 15, 27-29, 38
1.6	Collect documentation, reports and feedback for all Network training events and schools and publish on website.	1	ULUND	R	PU/PDE	3,9,23,36,48
2.3	Presentation from fellows at Network events and final conference	2,7	UNIGE	PDE	PU	9, 23, 36, 48
7.1	Data challenge website online and open to the public	7	CERN	PDE	PU	14
7.2	Report on commmunication, outreach and SMARTHEP Masterclass/WWDD (yearly).	7	CERN	R	PU	23, 36, 48
7.3	Report on presentation of results at international conferences;	7	CERN	PDE/R	PU	24, 36, 48
2.4	ESRs with three-year PhD projects receive a degree	2	UNIGE	ADM	CO	44
7.4	Report on publication of ESR results in peer-reviewed journals	7	CERN	PDE/R	PU	23, 36, 48

R: Report, ADM: Administrative; PDE: dissemin.; PU: Public; CO: Confidential, restricted to consortium and Commission services

B3.1.3 List of major milestones

The following table lists the major milestones of SMARTHEP. The milestones and their schedule have been designed so that the work on the WPs can take place as planned in Sec. 1.1.1, starting with a review of the state-of-the-art, proceeding with the work on software and tools and concluding with exploitation, dissemination and communication. Their timing allows both HEP and industry deliverables to have a check-point with sufficient margin for discussion and corrective measures if needed, in advance of the deliverable dates.

No.	Title	Related WP(s)	Lead Beneficiary	Due (mo.)	Means of verification
1	Grant agreement signed	1	ULUND	1	All beneficiaries have signed the grant agreement on the portal
2	Kick-off meeting	1	ULUND	2	The Kick-off meeting is held to launch the project
3	SB meetings	1	ULUND	3,9,23,48	Minutes of each SB meeting written and approved
4	School completion	2	UNIGE	12,17,31,40	Proceedings of schools available
5	Approval of all the PCDPs	1,2	ULUND	12	The PCDPs are approved by the Supervisory Board
6	Draft of SMARTHEP whitepapers on state of the art	3-6	CNRS, SU, TUDO, IBM	12,13,19,20	Whitepaper distributed to consortium and discussed in EB / on mailing list
7	Report on readiness of improved traffic/logistics apps, IoT and financial transaction tools after first secondments	4-6	SU, TUDO, IBM	23	Report given at the yearly meeting by ESRs and partners involved.
8	Readiness of ML and trigger algorithms for Run-3	3	CNRS	25	Algorithms developed by ESRs can be included in experiment software frameworks during commissioning of Run-3 LHC data taking
9	Readiness of hybrid architecture SW/HW for HEP	4	SU	33	Software for hybrid architectures can be tested using first Run-3 data
10	End of secondments	2	All	39	Fellows have concluded their secondments
11	Results presented at final conference	7	All	42	Public results of SMARTHEP presented in posters and talks by ESRs
12	Drafts of SMARTHEP whitepaper on ESR research	3-6	see 6.	34,41	Whitepaper distributed to consortium and discussed in EB before expected publication
13	Outreach activities concluded	7	All	48	Final public lectures and outreach on last Network-wide event
14	Conclusion meeting	1	All	48	ETN activities conclude with the celebration of the last Network-wide event and next steps are planned

B3.1.4 Overall activity plan

The activity plan of the whole network, including network events (Sec. 1.2.1), schools and milestones (Sec. 3.1.3) is shown in Fig. 4.

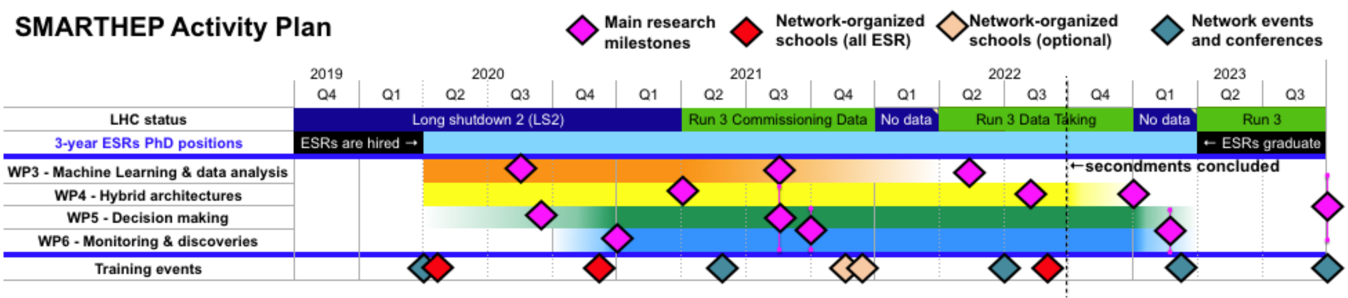

Figure 4: SMARTHEP activity plan

B3.1.5 Individual Research Projects

<table>
<tr><td>Fellow ESR1</td><td>Host UH</td><td>Ph.D. Yes</td><td>Start (mo.) 8</td><td>Duration (mo.) 36</td><td>Deliverables 3.1, 3.2, 3.3, 6.1, 6.2, 6.3, 6.4</td></tr>
<tr><td colspan="2">Work Package: WP3,6</td><td colspan="2">Doctoral programme: UH</td><td colspan="2">Title: ML and RTA for Higgs boson measurements and industry</td></tr>
<tr><td colspan="6">Objectives: Deep learning (DL) based algorithms which identify heavy objects (e.g. b quarks) in jets, e.g. DNNs, have been tested offline but not yet used in the trigger. As a first objective, ESR1 will understand the resource cost of these DNNs, improve them if necessary, and adapt them to be used in the trigger. This will significantly improve trigger performance, record more interesting data, and allow pure RTA-based data taking to exploit discriminating features previously inaccessible. The industrial secondment to Fleetmatics will have the second objective of adapting DL frameworks, e.g. Tensor Flow and Keras, in resource-constrained environments (e.g. embedded platforms), for real-time processing of images captured by mobile devices. This is an instrumental step towards mobile-phone in-vehicle edge computing applications for KKT. At KKT, ESR1 will be trained in theory and best practice of DL, and will return with expertise in RTA DL frameworks in constrained environments to improve the initial trigger selections towards their application in physics analysis. ESR1's third objective is to exploit DL in trigger algorithms to implement a generic RTA that caters to a variety of different physics analyses with hadronic activity and allows a much higher event rate. During the 5-month secondment to CERN ESR1 will be trained in both ML and RTA in the context of HEP, and work on the implementation and deployment of this general purpose scouting stream. ESR1 will validate this stream by measuring the frequent and well understood $Z \to b\,b$ process, and then apply it to a new measurement of the $H \to b\,b$ process.</td></tr>
<tr><td colspan="6">Expected Results: 1. Improve flavor/heavy object tagging at HLT in CMS (peer-reviewed paper). 2. Implement mobile ML frameworks for in-vehicle image processing at KKT. 3. Establish high-rate general purpose RTA stream for Run3. 4. Validate stream by measuring $Z \to b\,b$ cross section, then measure boosted $H \to b\,b$ production (peer-reviewed papers). ESR1 will receive a PhD in experimental HEP at University of Helsinki.</td></tr>
<tr><td colspan="6">Secondments: KKT, 4 months, Taccari, In-vehicle image processing in resource-constrained environment; CERN, 5 months, Pierini, commissioning generic data scouting analysis for Run 3 using DL.</td></tr>
</table>

<table>
<tr><td>Fellow ESR2</td><td>Host IBM</td><td>Ph.D. Yes</td><td>Start (mo.) 8</td><td>Duration (mo.) 36</td><td>Deliverables 3.1,3.3,3.2,6.1, 6.3, 6.3</td></tr>
<tr><td colspan="2">Work Package: WP3,6</td><td colspan="2">Doctoral programme: ULUND</td><td colspan="2">Title: Novelty detection for industry and ATLAS searches</td></tr>
<tr><td colspan="6">Objectives: Unlike SM particles, dark matter can have a longer lifetime than known particles and for example decay in the ATLAS calorimeter without leaving a trace in previous detector volumes. In this case, calorimeter "noise bursts" have a very similar signature and while relatively infrequent occur in 0.015% of the events, far more frequent than dark matter. ATLAS cannot currently distinguish the two in real-time. Anomaly detection is an ML technique concerned with the detection of ultra-rare "anomalous" events which do not follow part of the "normal" pattern of input samples and where little is known about the distribution of these anomalies. A wealth of different techniques exist, and this project will use existing ML methods to detect novelties, while extending and/or specializing the methods where appropriate to enable their use in RTA. Of particular interest to IBM is the possibility to combine ML with powerful general mathematical programming solvers, such as the IBM CPLEX commercial product. The first objective of ESR2 is to develop an algorithm that runs on existing ATLAS data and datasets provided by IBM, and discriminates between an anomalous signal and background, using CPLEX. Subsequently, this algorithm will be compared to the open source version, and the latter implemented as an anomaly detection system running in the ATLAS trigger, to be implemented in Run-3. The physics side of the project will be co-supervised by Doglioni (ULUND). While benefiting from her expertise in searches for new physics, this work will represent a significant step beyond the ATLAS physics program, since it targets final states never explored before. The 6-months CERN secondment under the supervision of Boveia (OSU) will ensure that this work is well integrated in the trigger system of the ATLAS experiment, and extend it to different detector signatures.</td></tr>
<tr><td colspan="6">Expected Results: 1. Develop combination of ML and programming solvers for novelty detection (peer-reviewed paper). 2. Extend improved novelty detection algorithms to RTA in industry and HEP (conference paper). 3. Apply to dark matter searches in ATLAS (peer-reviewed paper). ESR2 will receive a PhD in experimental HEP at Lund University.</td></tr>
<tr><td colspan="6">Secondments: CERN, 6 months, Boveia. Application of anomaly detection techniques to ATLAS data.</td></tr>
</table>

<table>
<tr><td>Fellow ESR3</td><td>Host IBM</td><td>Ph.D. Yes</td><td>Start (mo.) 8</td><td>Duration (mo.) 36</td><td>Deliverables 3.1,3.2, 3.3, 6.3, 6.1, 6.3</td></tr>
<tr><td colspan="2">Work Package: WP3,6</td><td colspan="2">Doctoral programme: SU</td><td colspan="2">Title: Real-time rule induction in fraud detection and HEP</td></tr>
<tr><td colspan="6">Objectives: An example of real time decision making that combines analytics and knowledge based models is fraud detection, notably in banking. Payment platforms detect in real time fraudulent transactions by combining recognition of human created patterns articulated on their banking symbolic knowledge model, and predictive models run to discover emerging fraud patterns by detecting new trends and anomalies from the data. Organizations usually manually inject or transform their machine learnt patterns into symbolic models for interpretability, reproducibility and transparency. The goal of this PhD thesis is instead to automate the learning of a symbolic model. In this thesis the student will work on new combinations of statistical and knowledge based models for a better decision automation in fraud detection and in high energy physics, for the recognition of human-created (fraud) and non-human-created (particle collision) patterns. While machine learning has been highly popular during the last years, their black box approach raises interpretability and explainability challenges. On the other hand, symbolic models, including rules, have been successful in making decisions more interpretable. Nevertheless they require to capture an existing knowledge or theory. In the context of real time decision automation ESR3 will test the proposed numeric to symbolic model inferences to detect anomalies, patterns and anti-patterns, combining the efficiency of the numerical machine learning and the explainability of the symbolic approach. They will inject theory (knowledge from the Standard Model in physics, fraud detection patterns in financial transactions) and combine it with predictive models to classify observations and add an interpretability layer. ESR3 will explore different angles in how we intend to combine numerical and knowledge models. As a first objective, they will induce a symbolic decision logic from numeric data, leveraging best breed predictive algorithms on data and infer symbolic rules from the predictive models. Then ESR3 will propose new algorithms to learn from the data and induce rules. The goal is to better capture and abstract the rules, while minimizing resource consumption. The third objective of the thesis is to apply these techniques to the real-time analysis of LHCb data, in a secondment with Gligorov from CNRS, to prototype a trigger that is able to distinguish new physics processes from known physics. Instead of deciding a priori how to accept or discard events, the trigger would "learn" the characteristics of the known Standard Model events from this dataset, and select events that do not fall into this category for further inspection. ESR3 will additionally be seconded to ULIEGE, to explore advances in autoregressive generative models for the purpose of searching for new physics or detecting anomalies, in the case of time-ordered data (e.g. to spot recurrent detector malfunctionings or repeated frauds). This family of neural networks has recently demonstrated its capability of learning highly structured data, such as images or sound, at a high level of fidelity, and offers access to the evaluation of the likelihood directly enabling the detection of unlikely events.</td></tr>
<tr><td colspan="6">Expected Results: 1. Combine numerical and predictive models for anomaly detection (peer-reviewed paper). 2. Design new algorithms for fraud detection (peer-reviewed paper, commercial software) 3. Apply algorithms to LHCb trigger at CNRS (conference paper). 4. Investigate time-ordered series at ULIEGE (conference paper). ESR1 will receive a PhD in computer science at Sorbonne University .</td></tr>
<tr><td colspan="6">Secondments: CNRS, 3 months, Gligorov. Application to LHCb trigger. ULIEGE, 4 months, Louppe. Investigation of autoregressive generative models.</td></tr>
</table>

Fellow ESR4	**Host** KKT	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 3.1,3.3, 4.1, 4.1, 4.4, 6.4
Work Package: WP3,6		**Doctoral programme:** UNIBO		**Title: RTA through computer vision on dashcams**	

Objectives: ESR4 will be trained in state-of-the-art computer vision algorithms based on deep learning and will study how to adapt and specialize them for RTA of videos collected by dashcams (camera on vehicle) installed on KKT customers. In particular, two classical problems in modern computer vision are depth from monocular images and semantic segmentation. Depth from monocular images is concerned with creating algorithms that can estimate the metric distance of objects from the camera using only one image, a problem traditionally solved with two cameras (stereo setup). Semantic segmentation algorithms assign a label among a predefined set of classes to every pixel of an input image (e.g. road, grass, pavement, pedestrian, etc...). The state-of-the-art approaches in both problems are nowadays based on highly-specialized deep learning pipelines, which require powerful GPUs to achieve real-time performance. Therefore, they are not suitable for deployment in smart dashcams to analyze video streams according to an edge-computing paradigm. The first objective of ESR4 will be the development and field testing of algorithms to compute depth from monocular images and semantic segmentation suitable for resource-constrained platforms, like the nVidia Jetson TX2, Ambarella CV2AQ, or Raspberry Pi 3 by using frameworks like Tensorflow Lite. ESR4 will propose simplifications of existing models to make them meet the real-time constraint on embedded devices and will test their accuracy on publicly available datasets as well as internal real customer data. The second objective of the project will be to investigate if and how hybrid platforms, like boards equipped with GPUs and FPGAs, can offer a different, more cost-effective solution to the problem of RTA of video streams in embedded platforms. A secondment in LIP6 in SU will train ESR4 with the background needed in heterogeneous computing architectures and the optimal resource allocation that ESR9 will develop expertise on, which will then be applied to the domain of this project. During the 2-months secondment at ULUND, ESR4 will use simulated datasets within the Open Data Project to deliver a starting point for deep learning techniques in HEP triggers, where the initial testing ground will be using energy deposits left in the detector by hadronic jets as images. This work will also have connections with ESR1's project for identification of jets from heavy quarks.

Expected Results: 1. Development and testing of algorithms for resource-constrained platforms (conference paper). 2. Investigation of hybrid platforms, together with LIP6 (peer-reviewed paper). 3. Proof-of-concept of computer vision algorithms on jet images (conference paper) ESR4 will receive a PhD in Computer Science and Engineering from University of Bologna.

Secondments: SU, 3 months, Lacassagne. Heterogeneous computing for video streams in embedded platforms. ULUND, 2 months, Doglioni. Sample toolkit for classification of hadronic jets.

Fellow ESR5	**Host** UHEI	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 4.1, 4.4, 4.4, 5.1, 5.5, 6.4
Work Package: WP4,5		**Doctoral programme:** UHEI		**Title: Real-time noise reduction new physics searches**	

Objectives: ESR5 will be trained in real-time methods for noise reduction which they will apply to the analysis of ATLAS data. This project will provide ESR5 with an expert-level understanding of the installation, calibration and operation of the large-scale high-energy physics experiments trigger systems, knowledge in the statistical analysis of experimental data in searches for new physics phenomena. ESR5 will learn how to code the programmable hardware for the high-speed real-time data processing of the detectors measuring particle energy in ATLAS (calorimeters). The first objective of ESR5 is the development and validation of the pileup noise reduction algorithms in the ATLAS calorimeter trigger system. This information will be used in real-time and cross-calibrated using the tracking information from the FTK, as a second objective. This will be aided by a secondment at CERN, supervised by FTK experts. The third objective of the project is the suppression of the known SM backgrounds in a search for Dark Matter particles using an angular analysis that depends critically on the noise reduction and calibration ESR5 developed. The use of a RTA will allow the subtraction of noise and identification of interesting events at the earliest possible stage. This will benefit from a secondment at ULUND under Doglioni's supervision, given her and her institute's expertise on this kind of searches, and from a secondment at XIMANTIS to work on hybrid networks (e.g. Convolutional Neural Network followed by Long Short-Term Memory recurrent networks) which can capture features in different dimensionalities (e.g. space and time for traffic monitoring and forecasting, position and timing of noise in the detector).

Expected Results: 1. Develop algorithms for better noise suppression in calorimeter hardware (peer-reviewed paper). 2. Deploy in ATLAS trigger with the aid of tracker information from FTK (software, conference paper). 3. Apply technique to angular searches (peer-reviewed paper). ESR5 will receive a PhD degree in HEP at Heidelberg University.

Secondments: UO, INFN (at CERN), 3 months, Strom, Boveia. Operations of FTK for trigger-level pileup suppression algorithm. ULUND, 2 months, Doglioni, application of noise reduction techniques to angular searches. XIMANTIS, 3 months, Sopasakis. Enhancement of AI modelling algorithms for traffic prediction, using hybrid networks.

Fellow ESR6	**Host** CNRS	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 3.1, 4.1, 4.4, 4.4, 6.2, 6.3, 6.5
Work Package: WP3,4,6		**Doctoral programme:** Sorbonne U.		**Title: Real-time trajectory reconstruction in ATLAS**	

Objectives: The hardware tracking processors FTK (Run-3) and HTT (Run-4) allow ATLAS to find particle trajectories in real time. Their performances such as efficiency and resolution are not determined just by the hardware, but also by a learned database of patterns (pattern banks) and geometrical constants. ESR6's first objective will be to deploy statistical and computing techniques, e.g. Principal Component Analysis and Graph Clustering, in order to produce new pattern banks for FTK and evaluate their impact on raw performances and physics analyses. ESR6's second objective will be to extend the application of FTK-like algorithms outside of HEP, by providing flexible and powerful tools to train datasets. These inter-sector toolkits will be developed during the secondment at KKT, and tested at INFN, whose geographical proximity allows for a seamless integration of the work. At KKT, ESR6 will be trained in the continuous online learning of ML models based on streams of labelled data, and in both the Apache Spark and the Amazon Web Services infrastructures for massively parallel computations. ESR6 will develop an online learning tool to continuously process real-time GPS data from Fleetmatics customers, providing customers with smart insights, improving customer experience, and thus increasing engagement with KKT products. The acquired ML and online processing skills will feed back into the main research project and will be applied to the FTK working dataset production. ESR6's third objective will be to investigate in detail the impact of new training on selected physics cases. In particular FTK tracks will be used to improve jet reconstruction and calibration, namely for the suppression of pile-up jets and the track-based components of the global sequential calibration, in collaboration with ESR5. This will enhance the sensitivity of fully RTA-based searches for dark matter mediators in the dijet mass distribution in Run 3 data.

Expected Results: 1. Deploy new statistical and computing techniques for FTK/HTT pattern banks (peer-reviewed paper). 2. Use of algorithms and implementation in Apache Spark / Amazon Web Services for online learning tool of GPS data, cross-talk with HEP (KKT/ INFN). 3. Use of FTK tracks in improved RTA-based search for dark matter mediators. ESR6 will receive a PhD in experimental HEP at SU.

Secondments: KKT, 4 months, Sambo, development of an online learning tool for GPS data processing. INFN, 3 months, Roda and Annovi, use of toolkits for creation of FTK pattern banks.

Fellow ESR7	**Host** UNIGE	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 3.2, 3.2, 3.3, 4.4, 4.4, 5.5, 6.2
Work Package: WP3,4,5,6		**Doctoral programme:** UNIGE		**Title: ML pattern recognition for exotic physics and industry**	

Objectives: Real time particle track reconstruction in LHC experiments is particularly difficult because of extremely busy detector images created by the multiple proton interactions (pile-up) in each bunch collision. This challenge will only increase in the future LHC upgrade. ESR7's first objective is to develop ML-based track reconstruction in ATLAS, as a replacement to algorithms that are too slow to be used in the trigger. ESR7 will be trained in track reconstruction, modern ML techniques, and general pattern recognition tools. This expertise will be crucial for ESR7's second objective and secondment with LIGHTBOX, in which ESR7 will utilise ML techniques to collect and analyze data from IoT sensors in industrial production chains in order to forecast in real-time when the machinery needs intervention. ESR7's third objective will be to evaluate GPUs for track reconstruction at the higher pile-up conditions of the LHC upgrades. ESR7 will compare GPU-optimized ML reconstruction to CPU-based reconstruction and dedicated hardware (e.g. FPGA) solutions proposed for ATLAS. ESR7 will receive dedicated co-supervision in optimizing algorithms for modern computing architectures and GPUs by by Martinez-Santos. ESR7's will then apply their knowledge to a novel real-time displaced vertex selection in ATLAS, one of the most promising and experimentally challenging NP signatures. ESR7's search for exotic long-lived (LLP) signatures with this selection will be a significant step beyond ATLAS's current capabilities, answer crucial questions for the future of ATLAS and open new avenues in the searches for new physics.

Expected Results: 1. ML-based trigger-level tracking reconstruction software in ATLAS (peer-reviewed paper). 2. Intervention forecasting algorithms for IoT-ready plants in LIGHTBOX. 3. Comparative assessment of GPU-based tracking. 4. Application to long-lived particle search (peer-reviewed paper). ESR7 will receive a PhD in experimental HEP at Université de Genève.

Secondments: Lightbox, 6 months, Catastini, ML for real-time industrial sensor data acquisition and analysis. USC, 2 months, Martinez-Santos, cross-experiment GPU optimization using open data.

Fellow ESR8	**Host** CERN	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** , 4.1, 4.2, 4.4, 4.4, 5.1, 5.5,
Work Package: WP4,5		**Doctoral programme:** UNIGE		**Title: Efficient RTA in ATLAS using multi-threaded processing**	

Objectives: Multithreaded (MT) programming is crucial to make best use of today's parallel computing architectures, but until recently most HEP code was unable to run MT. Because of associated overheads, MT is particularly challenging for RTA. ESR8's first objective will be to implement new monitoring within the ATLAS real-time code, measure algorithm scheduling and performance as well as the overhead of MT, and identify improvements that maximize MT performance. This will be done in synergy with ESR10 and ESR11's benchmarking work, and integrated in the ATLAS real-time software. Working on this objective will result in ESR8 becoming trained in advanced techniques of developing and evaluating code for highly parallel architectures. This will be crucial for their secondment to Lightbox, where they will devise an optimal parallelization of algorithms for investment strategies, trading infrastructures and integrated business processes. ESR8's second objective will be to be trained in these commercial tasks and then produce a commercial framework with figures of merit for their real-time optimization. ESR8's will also be seconded to CNRS and Sorbonne U. , whose physical proximity enables ESR8 to benefit from the expertise in MT and parallelization more generally of both. ESR8 will receive further MT and parallelization training and apply their knowledge on MT optimization to the creation of pattern banks for track triggers, working closely with ESR6. ESR8's third objective is to use the gained insights and knowledge to implement new RTA capabilities in the ATLAS trigger for LLP signatures, including dedicated pattern recognition algorithms for charged particles decaying in the middle of the detector. This will increase the trigger acceptance for such particles in ATLAS Run 3 data, so that ESR8 can lead a search with the first data, in collaboration with ESR7 and benefiting from the supervision of Sfyrla, as a fourth objective.

Expected Results: 1. Implementation and monitoring of MT algorithms in ATLAS trigger (peer-reviewed papers). 2. Commercial toolkit for the real-time optimization of parallel/sequential complex tasks in LIGHTBOX. 3. Use of MT in FTK pattern bank creation (CNRS, Sorbonne U.). 4. New trigger algorithms for LLP and search (peer-reviewed papers). ESR8 will receive a PhD in experimental HEP at UniGe.

Secondments: Lightbox 4 months, Catastini, improved efficiency for complex tasks by real-time decision of sequential/parallel processing. CNRS and Sorbonne U. , 5 months, Crescioli and Lacassagne, optimization of parallel code for FTK pattern bank creation.

Fellow ESR9	**Host** SU	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 3.1, 3.3, 4.1, 4.4, 4.4
Work Package: WP3,4		**Doctoral programme:** SU		**Title: RTA on heterogeneous computing architectures**	

Objectives: The goal of this project is to, in partnership with a number of computing hardware companies, train ESR9 in programming for heterogeneous computing architectures, and in simultaneously optimizing data formats and processing techniques to enable CPUs, GPUs, FPGAs, and hybrids to work together to solve problems which none of these technologies could solve on their own. The first objective of ESR9 will be to design a novel ML method for optimizing heterogeneous computing architectures, and deploy it in the context of the specific requirements of real-time data processing in the LHC collaborations. This will use ATLAS and LHCb as main examples to exploit synergies with ESR6 and ESR3 who are also based in the Paris area. This method takes as input the cost of the various computing architectures under consideration, builds and emulates test processing systems, tests them using the task being optimized for and finds the most cost-effective one. The second objective will be to apply this optimization code to the specific problems of the LHCb and ATLAS real-time data-processing architectures. This work will be done in collaboration with ESR10 and ESR11, as they will prepare a software toolkit that can be used to benchmark these processes. A five-month secondment at CERN will allow for training in ML methods and interaction with the LHCb and ATLAS collaborators involved in the optimization of their respective real-time analysis systems, under the supervision of Petersen and Couturier. Even though ESR9 does not have an industrial secondment, while seconded at CERN, ESR9 will be able to benefit from the mentoring of Catastini from LIGHTBOX, to understand the market potential of such advancements of hybrid architectures.

Expected Results: 1. ML-based optimization of heterogeneous computing architectures (peer-reviewed paper, software package). 2. Application of optimization code to LHCb and ATLAS triggers (peer-reviewed paper). ESR9 will receive a PhD in experimental HEP at Sorbonne University .

Secondments: CERN, 5 months, Petersen and Couturier, optimization of trigger systems of LHCb and ATLAS.

Fellow ESR10	**Host** NIKHEF	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 3.2, 3.3, 5.1, 5.3, 5.5, 6.4
Work Package: WP3,6		**Doctoral programme:** RU		**Title: Optimization of RTA resources and ATLAS LFV search**	
Objectives: Instead of focusing on specific physics processes to choose what is "interesting", ESR10 will investigate a more generic approach - what are the bottlenecks in our trigger algorithms, what prevents us to record events we want and how to do more interesting physics with the same or even fewer resources. ESR10's first objective will be to simplify, streamline and optimize the process of testing and benchmarking ATLAS triggers, by creating software tools that will analyze the performance of each trigger line, the resources used by each line, and their commonalities. ESR10 will also study how to reduce the consumption of resources by trigger lines using ML algorithms. A secondment at CERN under the supervision of physicists from the UO will prove that these tools are useful within the ATLAS trigger, testing the performance of the algorithms developed by ESR10 on chains that make use of FTK. As the tools developed by ESR10 will be experiment-independent, they will be portable and useful for applications for other LHC and future experiments. Specifically, we will design the tools together with LHCb trigger colleagues (ESR11), and for different computing architectures (ESR9). The secondment at XIMANTIS will allow ESR10 to use test a specific kind of algorithms, Attention networks, to focus both traffic prediction and resource consumption algorithms to a limited number of features. ESR10 will monitor the performance of these algorithms compared to the existing ones in the app, in an effort to make the developed tools fully environment independent, and provide a set of improvements for the app as its second objective. ESR10 will then test whether this kind of networks can help increasing the efficiency of trigger lines. ESR10's third objective will be to search for LFV in the $\tau \to 3\mu$ process with ATLAS Run 3 data, where the optimization of the analysis chain resource consumption is critical and will use previously developed tools.					
Expected Results:1. With ESR11, determine resources and bottlenecks in trigger systems, and identify how to optimize them (toolkit and peer-reviewed paper). 2. Establish optimal algorithms among new ML techniques for Ximantis traffic app. 3. Implement optimization for FTK triggers chains (with UO at CERN) and share code with LHCb (with CERN). 4. Search for LFV with ATLAS Run 3 data with optimized triggers (peer-reviewed paper). ESR10 will receive a PhD in experimental HEP at Radboud University, Nijmegen.					
Secondments: UO, CERN (at CERN), 5 months, Strom, Couturier. Benchmarking of algorithms for FTK in the ATLAS trigger and for LHCb. XIMANTIS, 4 months, Sopasakis. Optimization of the Ximantis app and testing of Attention networks.					

Fellow ESR11	**Host** NIKHEF	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 3.2, 3.3, 5.1, 5.3, 5.5, 5.4,
Work Package: WP3,5		**Doctoral programme:** VU		**Title: Optimization of RTA resources and LHCb LFV search**	
Objectives: ESR11 will collaborate closely with ESR10 on developing the toolkit for the optimization of the real-time trigger system of the main HEP experiments, and receive training in the same basic methods and techniques. ESR11 will focus on LHCb. Compared to ATLAS, LHCb generates an order of magnitude less data per collision at an order of magnitude larger rate. The requirement that the toolkit must work optimally for both experiments implies that it must be sufficiently generic and adaptable. As a result, it will also have applications beyond these two experiments. The workload for the creation of this toolkit will be split between ESR10 and ESR11, as both students will be working in synergy on the same topics, and will benefit from the supervision of both Igonkina and Raven. In addition both ESRs will adapt and optimize the toolkits for their respective experiments. A two-month secondment to Dortmund will allow ESR11 to receive expert training in the design of LFV analyses from Albrecht, and benefit from synergy between the work of Albrecht's group, the work of Albrecht's ESRs and the work of ESR14 who will be also seconded there. This will happen in tandem with a secondment at POINT8 where optimization and benchmarking techniques will be used for monitoring and decision-making in German industrial transport. Like ESR10, ESR11 will also be seconded to CERN towards the end of their PhD, to implement the benchmarking techniques on ATLAS and LHCb online systems. The work of ESR11 will be applied to the search for LFV in the $\tau \to \mu\gamma$ decay in LHCb, which currently does not have a dedicated trigger chain at LHCb. The high data rate of the upgraded LHCb experiment, in combination with the optimization of the trigger algorithms will allow to collect this kind of events and study a process that was previously thought to require dedicated experiments.					
Expected Results: 1. With ESR11, determine resources and bottlenecks in trigger systems, and identify how to optimize them (toolkit and peer-reviewed paper). 2. Apply techniques in analysis of public transport and logistics at POINT8. 3. Apply techniques to design an optimized trigger chain for LFV in the $\tau \to \mu\gamma$ decay in LHCb (peer-reviewed paper). ESR11 will receive a PhD in experimental HEP at VU University Amsterdam.					
Secondments: TUDO and POINT8, 5 months, Albrecht, Design of LFV analysis and RTA optimization of algorithms for public transport optimization. CERN, 2 months, Petersen and Couturier, application of benchmarking algorithms to ATLAS and LHCb.					

Fellow ESR12	**Host** TUDO	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 3.1, 3.3, 3.2, 5.1, 5.3, 5.5, 5.2
Work Package: WP3,5		**Doctoral programme:** TUDO		**Title: Real-time ML for LFV in unflavoured meson decays**	
Objectives: ESR12 will work closely together with ESR13, and search for LFV in neutral meson decays. The precision of tests by older experiments in LFV in neutral meson decays, in particular for decays of ϕ, J/ψ and $Y(1S)$ mesons, could be improved by orders of magnitude in the LHCb upgrade. This will need a fully RTA-based analysis, because of the difficult separation between these decays and the overwhelming background. ESR12's first objective will develop ML-based real time selections for the unflavoured ϕ, J/ψ and $Y(1S)$ mesons decaying into the different lepton species. ESR12 will be trained in optimizing ML selection algorithms for RTA applications, benefiting from the existing expertise of the ERC StG group of J. Albrecht. ESR12 will be seconded to XIMANTIS, where they will receive training in current best practice in ML and AI and gain hands-on experience optimizing ML algorithms for traffic prediction applications. This additional industrial experience and training will allow ESR12 to gain perspective on what kinds of methods are best suited to different types of real-time problems. A secondment at USC followed by a secondment at CERN will allow ESR12 to deploy and commission the RTA techniques for LFV decays of light and unflavoured particles in the LHCb trigger as a third objective. ESR12's fourth objective will be to search for LFV in the decays of unflavoured mesons using Run 3 LHCb data.					
Expected Results: 1. Study of ML-based RTA for neutral meson decays. 2. Enhance existing AI-based traffic modelling at XIMANTIS. 3. Deploy and commission trigger in LHCb (peer-reviewed paper). 4. Search for LFV in decays of unflavoured mesons. ESR12 will receive a PhD in experimental HEP at Dortmund.					
Secondments: USC, 2 months, Martinez Santos, physics of LFV. CERN, 2 months, Matev, development and implementation of LFV trigger selections. XIMANTIS, 4 months, Sopasakis, enhancement of AI modelling algorithms for traffic prediction, using Convolutional Neural Networks (CNNs).					

Fellow ESR13	**Host** TUDO	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 3.1, 3.3, 3.2, 5.1, 5.5, 5.4
Work Package: WP3,5		**Doctoral programme:** TUDO		**Title: Global event triggering in LHCb**	
Objectives: HEP triggers first reconstruct objects (e.g. jets or tracks), and then perform a selection on these objects. Event triggers select based on overall (global) event properties, avoid time-intensive reconstruction, and permit searching for new physics in areas where traditional object-based selections are too slow. ESR13 will receive training in DL of complex systems, as well as other modern ML and AI methods, and apply these methods to global analyses of HEP collisions in RTA. ESR13 will focus on DL using recurrent neural networks, which have shown great success in speech recognition or translation, and use them to further improve global analysis tools. ESR13's first objective will be to analyze the overall pattern of detector hits to identify events enhanced in interesting physical processes, building on promising initial studies in Albrecht's ERC StG. As a first objective, ESR13 will design a global trigger selection and benchmark its performance against a traditional object-based approach. This will be done in collaboration with ESR11, within a secondment at NIKHEF. The experience in global event trigger selections will be used to collaborate with IBM's ESR3, comparing a symbolic approach with a purely stochastic ML one. With the secondment at POINT8, ESR13 will work on a real-time data-analysis project in German industry that requires a global analysis. As a second objective, the ESR will analyze client-provided data on real-time traffic prediction for transport companies, applying similar techniques as in HEP. Once back in Dortmund, ESR13's final objective will be to apply the developed event triggers to the analysis of LFU violation in semileptonically decaying beauty mesons, which are abundant enough to allow the event trigger to be benchmarked against more traditional approaches.					
Expected Results: 1. Global event-based RTA using recurrent NN and other ML algorithms (conference paper). 2. Implementation of global RTA to transport in POINT8. 3. Application of global event-based RTA and benchmarking against traditional approach (peer-reviewed paper) ESR13 will receive a PhD in experimental HEP at Dortmund.					
Secondments: NIKHEF, 4 months, Raven. Optimization of global event triggers. POINT8, 4 months, Dungs. Analysis of logistics data with RTA methods.					

Fellow ESR14	**Host** UHEI	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 3.1,3.3, 4.4, 3.2,6.1, 6.2, 6.3
Work Package: WP3,4,6		**Doctoral programme:** UHEI		**Title: RTA to search for Dark Photons in LHCb**	
Objectives: ESR14 will develop an advanced fully RTA-based analysis to efficiently select decays involving displaced di-electron vertices with low invariant mass. ESR14 will use these advances to search for light dark photons with the LHCb detector focusing in a range of extremely low coupling so far unexplored by other experiments. The upgraded LHCb detector, which is planned to start data-taking in 2021, will allow for the first time to search for this kind of signature. It is the novel real-time detector readout in Run-3 that gives the possibility to use RTA to separate dark photon candidates from the overwhelming background online. The reconstruction of tracks and displaced vertices has to be carried out online even for low-momentum tracks. This challenging task will depend critically on the performance that can be achieved on modern computing architectures. ESR14 will be trained in the usage of highly parallel architectures for the LHCb online reconstruction and will profit from a secondment in USC on use of GPUs. ESR14 will develop advanced ML algorithms to optimise the electron classification and make it fast enough to be run in real time. This work will benefit from a secondment in TUDO under Albrecht's supervision, who will host both ESR14 and ESR11 and train them on how to design an online analysis of decays involving electrons with the help of his StG team. These advances will also allow ESR14 to develop a tool to efficiently retrieve online low momentum photons through their conversion in di-electron pairs, a challenging task that requires an RTA-based approach. Even though ESR14 will not have an industrial secondment, they will receive mentoring by Dungs from POINT8 (who was previously working in LHCb HEP and worked at Google) while in Dortmund and remotely, on matters of careers in industry starting from academia.					
Expected Results: 1. Development and testing of algorithms for RTA of electrons (conference paper). 2. Application to dark photon search with Run 3 data (peer-reviewed paper). ESR14 will receive a PhD in Experimental HEP from Heidelberg University.					
Secondments: TUDO, 3 months, Albrecht. Design of LFV physics analysis. USC, 3 months, Martinez Santos. Use of GPU in LHCb online reconstruction.					

Fellow ESR15	**Host** ULUND	**Ph.D.** Yes	**Start (mo.)** 8	**Duration (mo.)** 36	**Deliverables** 6.1,6.2,6.3
Work Package: WP6		**Doctoral programme:** ULUND		**Title: Real-time calibration and analysis of the ALICE TPC**	
Objectives: ESR15 will study and commission the ALICE Time Projection Chamber (TPC) detector. In 2019-2020 the LHC will be shut down to upgrade and prepare the experiments for Run 3. The goal of the ALICE upgrade is to be able to analyze the full rate of 50000 events per second, increasing the sensitivity for most measurements by between one and two orders of magnitude. The main goal on the detector side is the upgrade of the TPC with a Gas Electron Multiplier (GEM) readout that will allow continuous operation. This continuous readout requires a whole new software framework denoted O2 (online-offline), whose goal is to do full calibration and reconstruction in real time. ESR15's first objective will be to contribute to the development of the ALICEO2 framework and use this expertise in the analysis of the first data from Run 3, which will start in 2021. Particularly, due to the space charge build up fluctuations, the upgraded TPC will have to be calibrated every 5 milliseconds over a space volume of 90 cubic meters. This demands fast, effective and robust algorithms that ESR15 will have to develop, tune and benchmark. A 6-months secondment at CERN will allow the implementation of the real-time calibration algorithms in the ALICE software, as well as training and interaction with the core of the ALICE O2 development team. The second objective of ESR15 is the analysis of heavy-ion data, which will first be available at the end of 2021. This analysis will be a measurement of the nuclear modification factor, since this measurement will be very sensitive to the corrections and a perfect testing ground for the algorithms developed. Even though ESR15 will not have an industrial secondment, they will receive local mentoring by Sopasakis from XIMANTIS (also a lecturer in mathematics at ULUND) on matters of industry/academia interaction.					
Expected Results: 1. Contribute to development of the O2 algorithms for ALICE TPC reconstruction (peer-reviewed paper). 2. Apply algorithms to the measurement of the nuclear modification factor (peer-reviewed paper). ESR15 will receive a PhD in experimental HEP at Lund University.					
Secondments: CERN, 5 months, Shahoyan, implementation of algorithms in ALICE trigger software.					

B3.2 Appropriateness of the management structure and procedures

B3.2.1 Network organization and management structure

WP1 is dedicated to the management of SMARTHEP. As specified in the WP description, the Project Coordinator (PC) will be Doglioni from ULUND. She will benefit from the experience gained coordinating working groups in her collaboration, in the LHC-wide LHC Physics Center, as HEP Software Foundation convenor, as well as from her education that connected academia and industry. The PC will act as the bridge between the Network and the EU and will be in charge of the management structure of SMARTHEP, as well as of keeping an up-to-date overview of the network progress for the governing bodies of the network using the OmniPlan software and creating the infrastructure for network-wide communication as described in Sec. 2.4.1. The PC will also drive the signature of the **Grant Agreement** (GA) and **Consortium Agreement** (CA) between all the nodes. The process leading to this will define the final structure of the Network, the employment status of the ESRs, Intellectual Property Right aspects and training structure, and it will be steered by the PC and by the Work Package coordinators presented in Sec. 3.1.1.The WP coordinators will be responsible for the overall coherence and deliverables of their respective WP, the management of risks concerning their WP, and the reports of the progress in the WP they coordinate through the PCDPs. All WP coordinators including the PC and the Project Manager hired to support the PC will form the SMARTHEP management board and meet bi-monthly to discuss the network progress.

Alongside the management board, the Network will have two main discussion and decision-making bodies: an Executive Board (EB) and the Supervisory Board (SB). These are pictured in the following figures and described in detail in Sec. 3.2.2.

Official network communication, including decisions taken in EB and SC meetings, will be transmitted on the existing network e-mail list `smarthep-participants@cern.ch`. The archives of this mailing list will be available, except in matters requiring non-disclosure. In those cases, communication will happen between the interested parties according to the regulations of each of the participants involved. A "virtual corridor" for the network has already been created at CERN using Mattermost software for rapid, non-official communication and collaboration, allowing network members to create channels for discussion and effectively share documents, software, and announcements.

B3.2.2 Joint governing structure

The EB is the extended project management body of SMARTHEP. It is composed of the PC, coordinators of all WPs, a **student representative** elected from a board that the ESRs will be encouraged to form for their own discussion, three officers responsible for **recruitment, diversity and secondment** matters, and representatives from each **academic and industrial beneficiary and partner**. The EB is expected to meet via teleconference regularly (e.g. every quarter), as well as on demand. It is a forum where all WP coordinators and contact persons report on their progress, results or other issues relevant for the Network. Action items arising from these reports might be: a communication strategy for a particular publication, the formation of a task force to deal with an identified risk or failure, a new avenue of research or a new product to be developed. These meetings will be chaired by the PC and minuted with listed action items. Draft minutes will be released shortly after the meeting, and final minutes within a week. Each EB meeting will follow up on outstanding action items. The EB will meet physically during the first days of Network-wide events, and prepare the agenda for the SB meetings. The EB will also serve as a point of contact with all SMARTHEP stakeholders, and provides support to the nodes organizing Network-wide events. Finally, the EB will be able to propose changes to the plans presented in this proposal to the SB.

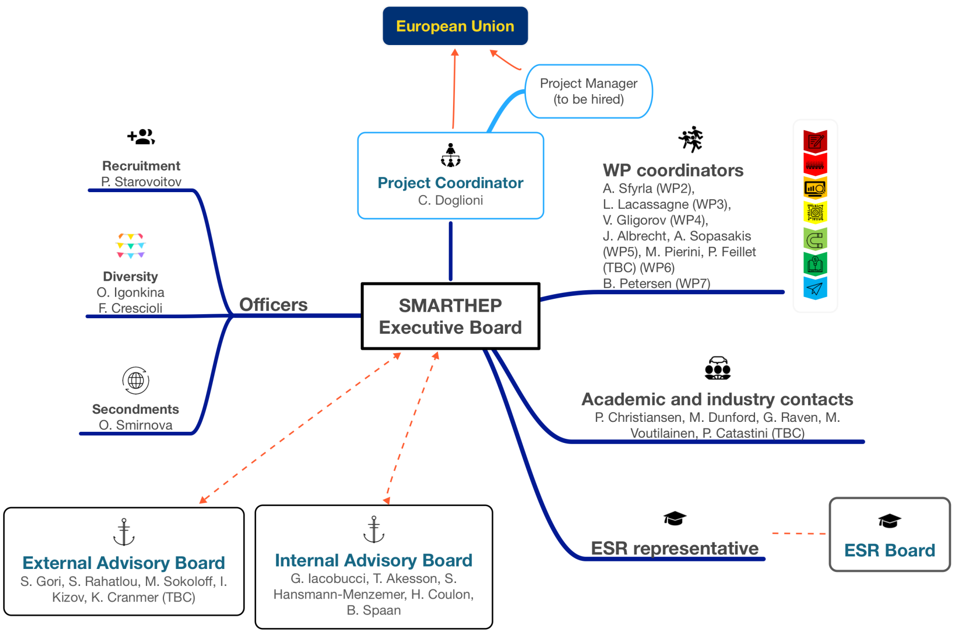

Figure 3: SMARTHEP Executive Board (EB)

of senior physicists that will help consolidating the managerial experience of the junior researchers in the network. The Internal Advisory Board is composed of senior professors from the beneficiary institutes mentioned in Sec. 1.3.1, who have extensive experience in project management and coordination. The External Advisory Board is composed of experienced professors that are not employed in beneficiary or partner institutes, who will also participate in the training events. It includes Prof. Stefania Gori from UCSC, an expert in theoretical physics phenomenology, Prof. Mike Sokoloff from U. Cincinnati, one of the founding member of the HEP Software Foundation, Prof. Ivan Kisel from FIAS, who is one of the pioneers of many reconstruction methods for particle tracks in high energy physics, Prof. Shahram Rahatlou from University of Rome Sapienza, who has been the CMS Physics Coordinator and Prof. Kyle Cranmer from NYU's center for Data Science, an expert in statistics, machine learning and data preservation. The members of this board will act as external observers to the network and offer advice on the impact and dissemination activities as seen from the broader community. They will also help resolve conflicts between participants that may arise within the network. The External and Internal Advisory Board members will be invited to the SMARTHEP yearly meetings.

B3.2.3 Supervisory board

The SB drafts the CA and is the decision-making body of SMARTHEP in matter of training and project, by majority vote. It receives advice and action items from the EB. Each SMARTHEP node has a single vote in the SB. They are usually represented through their main contact person as specified in the List of Beneficiaries and List of Partner Organizations Tables. The student representative will be allowed to participate and vote in the SB, but the SB reserves the right to hold closed meetings in case of items that are critical for one or more ESR (e.g. failed progress, crisis management, scientific misconduct). The approval of the PCDPs will not include the ESR vote. The work plan presented in this proposal will have to be approved by the SB at the Kick-off meeting, to be held during month 2 of the project. The SB will also have to approve any measurement proposed by the EB and the PCDP of all the ESRs. The SB will monitor the progress of the Network, including the quality of training and supervision, and will monitor and take action on scientific misconduct as per the the The European Code of Conduct for Research Integrity[36]. Apart from the Kick-off meeting, the SB will meet yearly in the special Network-Wide events organised by SMARTHEP. These meetings will be chaired by the PC.

B3.2.4 Recruitment strategy

SMARTHEP will recruit in compliance with the "European Code of Conduct for the Recruitment of Researchers". SMARTHEP is committed to appointing the ESRs in a gender unbiased way and candidates will be selected regardless of their race, religion, sexual orientation or disability. We are aware that marginalized groups are significantly underrepresented among job applicants and will take proactive steps to advertise SMARTHEP ESR positions among them and highlight their inclusiveness, as discussed in a dedicated section (Sec. 3.2.8). SMARTHEP will have a common job advertisement template and a dedicated information webinar for prospective ESRs at the kick-off meeting. Work on the recruitment will start in advance of the CA signature so that all ESRs are recruited by month 8. Positions will be announced and described in detail on appropriate portals at the national, European, and international level, e.g. on the Euraxess portal, our web page and social media.

Once recruitment for each node is completed, the corresponding beneficiary main contact will inform the recruitment officer and the relevant WP coordinator(s) and bring the proposal to the EB meeting for approval. As part of this approval, the two best-ranked runners-up will be reviewed by the recruitment and diversity officers, as EB members not involved in the original recruitment. The goal is to verify that the selection process was unbiased with respect to gender, race, religion, sexual orientation or disability. In particular, the review will establish whether there was any unconscious bias on the part of the recruiting panel. Starovoitov (Heidelberg) will serve as **recruitment officer** with a seat in the EB overseeing the recruitment procedure.

B3.2.5 Progress monitoring and evaluation of individual projects

Several progress monitoring and evaluation plans are foreseen:

- A PCDP will be created for each ESR, and approved first by the supervisors and then by the WP coordinator and SB.
- Supervisors will report on ESR progress to the WP coordinator where the ESR is working, as well as in EB and SB meetings. Supervisors will also communicate back the feedback coming from the ESR to the SB.
- Each secondment responsible (partner or beneficiary) will write a short progress report of the ESR's work and achievements at the end of the secondment, monitored by the dedicated **secondment officer** Smirnova (ULUND). Secondments are the basis of lasting relationships between the ESR and the network partners maintained beyond SMARTHEP.
- ESRs will present their work in talks and/or posters in the yearly network events, as well as in international conferences when major milestones are reached. The ESR presentations and posters will be collected by the **WP7 responsible** Petersen (CERN) on the SMARTHEP website.
- Yearly appraisal meetings will be held by beneficiaries, and chaired by the senior members of the institute or university.
- Each ESR will write half-time and final reports, which will be approved by supervisors and relevant WP coordinators, comparing projects outcomes to the PCDP.

Each WP will organise teleconference meetings whose frequency will depend on the size and activities of the group, and will have joint meetings as appropriate. ESRs will also be encouraged to participate in relevant meetings held by the experimental collaborations and industrial partners.

B3.2.6 Risk management at consortium level

Upon its creation and regularly during its meetings, the SB is tasked with creating a detailed Risk Management Plan and bringing it to the attention and approval of the EB. A preliminary list of identified risks and proposed mitigation measures relating to the deliverables and milestones defined in Tab. 3.1.2 and 3.1.3 are listed the table below.

[36] https://ec.europa.eu/research/participants/data/ref/h2020/other/hi/h2020-ethics_code-of-conduct_en.pdf

Risk No.	Description of Risk	WP	Proposed mitigation measures
R1 (low)	Delays in recruitment	1	Announcement of vacancies earlier or soon after signature of grant agreement, awareness of visa issues in consortium, eight month contingency.
R2 (medium)	Biases in recruitment	1	Targeted advertisements to minority groups, "blind" application process up to interview (only research interests visible up to shortlisting process), cross-check of recruitment approval in EB
R3 (low)	Problems in training or outreach event organization	2, 7	Two beneficiaries will be assigned to organize a given training or outreach event, with the help of the WP responsibles, to have sufficient support and overlap between scientists.
R3 (low)	Underperforming ESR	2-7	Excellence-based recruitment, continuous monitoring, discussions and feedback, possibility of awarding intermediate degree with one of the beneficiaries (e.g. Licentiate degree)
R4 (medium)	Practical difficulties in ESR mobility	3-6	Design of academic and industrial secondments requiring maximum 2 long-term stays away from home node, dedicated secondment officer
R5 (medium)	Impossible to perform research/secondment (e.g. supervisor changes position or career)	3-6	Extra backup projects from each of the partners related to ESR ready during proposal writing stage, more than one PI per beneficiary institute.
R6 (low)	Delays in LHC schedule	5	Preparation of algorithms and physics searches using simulated data, prototyping research objectives on data already taken.
R7 (medium)	Poor performance of planned algorithm/infrastructure	3,4	Since some of the techniques used are of a high-risk high-gain nature, this is the most relevant risk for the ESR progress. We will review and downgrade specifications as necessary, with focus on retaining most useful and exploitable characteristics, document work done and reevaluate ESR PCDP by reassigning tasks within the network objectives.

Risk management will be monitored by each WP coordinator as well as by the diversity, recruitment and secondment officers, following guidelines defined by the SB at the Kick-off meeting and included in the CA. The WP and research projects are designed such that any potential issue in a single WP or deliverable will not affect the remaining deliverables. In the case that any major failure in the development of a project is detected, the relevant WP coordinators will inform the EB. Together with the ESR, the supervisor(s) and the WP coordinators, the EB will design the best follow-up strategy and will monitor its implementation and outcome.

B3.2.7 Intellectual Property Rights (IPR)

IPR Management during the project.For the success of the project it is essential that all parties agree on explicit rules concerning IP ownership, access rights to any Background and Results for the execution of the project and the protection of intellectual property rights (IPRs) and confidential information before the project starts. Therefore, such issues will be addressed in detail within the CA between all project parties. Additional industrial partners have granted SMARTHEP access to machines with architectures beyond the state-of-the-art for testing their algorithms, under the agreement that the publication will follow the public release of the machine and that the company producing a given machine is not named until then. These machines are hosted at CERN and will be accessible to SMARTHEP researchers.

Consortium Agreement. The purpose of the CA is to establish a legal framework for the project in order to provide clear regulations for issues within the consortium related to the work, IP-Ownership, Access Rights to Background and Results and any other matters of the consortium's interest.

Access Rights to Background and Results. For smooth execution of the project, in the CA the beneficiaries will grant each other and their affiliated companies, royalty-free Access Rights to Background and Results necessary for the execution of the project. This will allow the researchers the ability to execute the project to the best of their ability, without being hindered by administrative issues. The CA will define further details concerning the Access Rights for Exploitation to Background and Results.

IP Ownership. Results shall be owned by the project beneficiary carrying out the work leading to such Results. It will be agreed among all parties in the Consortium Agreement that the ownership on IP rights will be solely retained at the host beneficiary of the corresponding ESR, when this knowledge is generated during the execution of the project at the host beneficiary. For Results created during secondments, specific one-to-one IP rights agreements will be signed on a case-by-case basis. If any Results are created jointly by at least two parties and it is not possible to distinguish between the contributions of each of the parties, such Results, including inventions and all related patents/patent applications, will be jointly owned by the contributing parties. In order to further the competitiveness of the EU market, and to enhance exploitation of the Consortium Results, each contributing party shall have full own freedom of action to exploit the joint IP as it wishes, to further the goals of the consortium. To promote this effort the contributing party will have full own consideration regarding their use of such joint Results and will be able to exploit the joint IP without need to account in any way to the other joint contributor(s). Further details concerning jointly owned Results, joint inventions and joint patent applications will be addressed in the CA.

Transfer of Results. As Results are owned by the project parties carrying out the work leading to such Results, each project party shall have the right to transfer Results to their affiliated companies without prior notification to the other project parties, while always protecting and assuring the Access Rights of the other project parties. Such use of Results will encourage competitiveness of the EU market by creating broader uses of the Results.

Open Source. A central aim of this consortium is to provide benefit to the European community. Some of the project parties may be either using Open Source code in their deliverables or contributing their deliverables to the Open Source communities. Details concerning open source code use will be addressed in the CA.
Website and legacy. As coordinator node, Lund University will maintain the SMARTHEP website and access to the software beyond the end of the project, ensuring that GA signatories retain fair-use access to the deliverables.

B3.2.8 Gender and equal opportunities aspects

All members of SMARTHEP believe in fostering a research culture where everyone including marginalized groups feel encouraged to collaborate, and hold this particularly important when training new scientists. SMARTHEP will appoint two dedicated **equal opportunity officers**, Prof. O. Igonkina and Dr. F. Crescioli, to ensure that during the recruitment and throughout the duration of the action SMARTHEP remains a welcoming research environment and trains both ESRs and its more senior members in matters of inclusion.

The coordinator of SMARTHEP, as well as more than 30% of the beneficiary contacts, are female scientists[37], close to the Commission's goal of 40% for decision-making panels. To mention only examples of the involvements of the SMARTHEP members in increasing diversity in science, Gligorov plays a leading role in promoting diversity and equality within the LHCb collaboration, and authored a statistical analysis of the gender evolution of the LHCb collaboration since its formation, UNIGE, CNRS, NIKHEF have been members of the https://genera-project.com/index.php/consortium and ULUND is an active member of the NORNDIP network financed by Nordforsk. These networks will be a source of support and local training for the ESRs.

As equal opportunities are a priority for SMARTHEP, we will actively promote diversity when recruiting the ESRs, while still keeping high scientific standards throughout the process, according to the European Charter for the Recruitment of Researchers[38]. In addition to having members of selection panels with different expertise, gender and personality, SMARTHEP will seek to avoid biases in our recruitment process through a two-step recruitment policy, with an explicit focus on the avoidance of unconscious bias. The SMARTHEP recruitment process will clearly state the equal opportunities policies in all our recruitment documents, and our recruitment advertising will prominently showcase e.g. the female supervisors and SB members of the consortium. SMARTHEP are aware of research indicating that minority job applicants self-select to great extent. SMARTHEP will therefore explicitly (in the job advert) and proactively encourage applications from female and minority candidates by targeting the recruitment advertisement to specific groups and clubs (e.g. associations of women in physics, LGBT at CERN, outreach groups targeting developing countries), and by highlighting the local diversity and inclusion committees and measures at the recruiting node in the advert itself. The recruitment progress will be reviewed at the Kick-off meeting of SMARTHEP, and if SMARTHEP find that fewer than 30% of applicants at that stage have been women and/or from minority groups, the SB will be tasked with developing immediate and explicit corrective measures in order to encourage more applications from female and minority students.

All the nodes of the Consortium have a clear commitment to inclusion and equal opportunities. SMARTHEP will make explicit the important roles that women and minority groups play in the Network through a section of its web page. Female ESRs and ESRs from minority groups will be active in dedicated outreach activities to show the importance of diversity in science, to act as role models and to demonstrate in practice the equality of conditions in a welcoming environment in science. All SMARTHEP members will be expected to use gender neutral language in network activities, pay particular attention to desired gender pronouns when referring to ESRs, and all SMARTHEP publications will use gender neutral language throughout. Lectures to avoid unconscious bias will be offered in our Transferable Skills program, so that our new generation of scientists is educated in the respect, inclusion and Equal Opportunities policy for all genders.

B3.3 Appropriateness of the infrastructure of the participating organizations

All nodes in the consortium have infrastructure appropriate to the tasks they are responsible for. ULUND's management is supported by its Research Services, which are appropriate for a large university. All ESRs will have access to the complete range of SMARTHEP expertise (Sec. 3.4) and infrastructure (part B2), and be provided with office space and a laptop for use during SMARTHEP. Beneficiaries and the secondment officer will also help with accommodation, visas and practical issues, making use of the EURAXESS Services for Relocation Assistance. Academic support granted to ESRs includes close supervision, mentoring by senior scientists and institutional academic literature subscriptions, as well all the facilities supporting research at the nodes. All institutes and research laboratories have software divisions with excellent computing capabilities and infrastructure, including rack-mounted servers and large data storage systems, locally and on the Worldwide Computing Grid.

As detailed for each node in part B2, SMARTHEP participants contribute with the specific facilities and equipment required for the successful completion of the WPs. In particular, partners hosting secondments involving GPUs and hybrid architectures host those on their premises and ESRs will be able to use them during the secondment and remotely on demand. Moreover, a

[37] This is significantly more than the average percentage of women in physics, see The gender gap in science: How long until women are equally represented?, PLoS Biol 16(4) e2004956 [38] `https://cdn2.euraxess.org/sites/default/files/brochures/am509774cee_en_e4.pdf`

number of industrial entities have agreed with Prof. Lacassagne that ESR9 will be able to make use of beyond-state-of-the-art machines to test algorithms developed throughout the duration of SMARTHEP, and that ESR9 will be able to disseminate results in open access journals as soon as the machines become public, within the course of their PhD project.

B3.4 Competences, experience and complementarity of the participating organisations and their commitment to the program

B3.4.1 Consortium composition and exploitation of participating organisations' complementarities

The academic beneficiaries and partners of SMARTHEP are the pioneers and experts of RTA in each of the LHC experiments. They have a strong background in the physics topics that are the focus of the network, as the network includes 5 ERC grantees and scientists who have covered and are covering positions of responsibilities within large international collaborations as detailed in Secs. 1.3.1 and part B2. Researchers in SMARTHEP have already been successfully collaborating within their experiments on complementary topics. Scientists in NIKHEF, CNRS, TUDO (LHCb), UO, OSU, ULUND, UHEI (ATLAS), UH (ATLAS) and ULUND (ALICE) lead the trigger and RTA techniques developments for each of the collaborations. Expertise in ML is given by POINT8, UNIGE, by LHCb researcher Ustyuzhanin from the Yandex School of Computing and CERN.

The synergy in the RTA physics program of SMARTHEP is cross-collaboration, with researchers from NIKHEF, CNRS and TUDO researching LFV/LFU in ATLAS and LHCb, researchers from UNIGE, CERN, UH, UHEI, UO seek new physics through Higgs and dark sector particles in ATLAS and CMS, while ULUND and UH precisely measure the Standard Model of particle physics in CMS and ALICE.

Software and machine-learning experts from POINT8, XIMANTIS, DELETEME-HIMT, CERN and IBM as well as experts in hybrid infrastructures in INFN, CNRS, KKT and LIGHTBOX provide the complementary tools to develop efficient solutions to both industry and research problems, and in the case of the industrial partners, practical implementations with a commercial value.

The status of CERN as an international collaboration that hosts all HEP experiments provides the ideal framework in which to exploit the research complementarity within the network and support its advancements beyond the state of the art. An example is the recently Internet-of-Things Openlab effort at CERN[39] whose organizers informally agreed to support this network through a half-day training and presentation day. The participation of all LHC collaborations to SMARTHEP and the explicit support of the HEP Software Foundation (see letter of support) offers a unique chance to form lasting collaborations and shape the future data taking strategies of LHC experiments and beyond, building from the results obtained in SMARTHEP.

B3.4.2 Commitment of beneficiaries and partner organisations to the program

While the majority of the SMARTHEP academic participants are already successfully collaborating in research projects within the LHC experiments, this is the first network that including all major LHC experiment that will consistently advance RTA through Machine Learning and hybrid architectures for modern particle physics experiments and industry alike. In-person meetings for the members of the SMARTHEP network have been taking place since May 2017[40]. A number of subsequent meetings have taken place, including the RAPID2018 workshop (Sec. 1.4.2), and culminated in a three-day workshop dedicated to the preparation and collaborative editing of this ITN [41]. SMARTHEP members have also secured funding for a 12 day RTA workshop at the Institut Pascal (Sec. 1.4.2). It is clear from the enthusiasm and participation of all members, as well as from the support offered to this network by ULUND, that all SMARTHEP participants are fully committed to this project.

High quality education of students in large-scale international projects is already an essential part of the mission of the universities and research labs in SMARTHEP. SMARTHEP will greatly strengthen and extend their capacities to carry out these activities, by dedicating a significant fraction of research and supervision time to SMARTHEP activities (see node descriptions in part B2). All industrial nodes are committed to SMARTHEP, to benefit from the experience in RTA of LHC experts, to foster connections to the academic world and to expand their activities. The industry parties have a strong motivation to participate in cutting edge developments and the subsequent transfer of the technologies to new commercial applications. SMARTHEP partners will not only contribute training to the network, but will also benefit in their own research and commercial objectives through the secondment projects of the ESRs.

The relations between the nodes, established starting from the preparation of the project will last even beyond SMARTHEP: to organize post-graduate recruitment events for companies in the network that specifically target SMARTHEP ESRs after the completion of their PhD and to continue the series of training events and offer them to the European graduate student community.

---

[39] First workshop of Internet-of-Things at CERN, `https://indico.cern.ch/event/669690/overview`, Nov 2017 [40] SMARTHEP kickoff meeting, supported by the Grace och Philip Sandbloms Fund [41] SMARTHEP writing meeting at CERN, supported by Lund Research Services