JENA White Paper on
European Federated Computing

The JENA Computing Initiative

March 23, 2025

The Joint ECFA-NuPECC-APPEC (JENA) Activities launched an initiative (JENA Computing) in 2023 to promote the increasing need for discussions on the strategy and implementation of European federated computing at future large-scale research facilities. In workshops and dedicated working groups on specific topics, expert groups from all relevant research areas were formed to compile an overview of existing strategies in the individual countries and communities. Here we present a summary of the resulting Working Group Reports, including the most important recommendations from these areas of computing. Furthermore, an additional chapter on sustainability in the field of computing is included. This version of the JENA White Paper on European Federated Computing serves as a basis for discussion at the JENA Seminar in April 2025 and as input to the European Strategy for Particle Physics - 2026 update (ESPPU), and may be revised thereafter.

Contents:

Executive Summary

Contributors:

JENA board representatives:
Paris Sphicas, Lidija Zivkovic (ECFA)
Eberhard Widmann, Sissy Körner (NuPECC)
Carlos Peña Garay, Julie Epas (APPEC)

Members of the JENA Computing Initiative:
Stefano Bagnasco, Concezio Bozzi, Simone Campana, Sascha Caron, Xavier Espinal, Andreas Haungs (chair), Alberto Iess, Andreas Ipp, Marc Labiche, Giovanni Lamanna, Paul Laycock, Marek Lewitowicz, Adrien Matta, Gonzalo Merino, Johan Messchendorp, Arnau Rios, Stefan Roiser, Thomas Schörner, Graeme A Stewart, Patrice Verdier

JENA White Paper on

European Federated Computing

The JENA Computing Initiative - March 23, 2025

Executive Summary

The **JENA Computing Initiative** assesses trends, needs, and strategies in European federated computing for particle, astroparticle, and nuclear physics at current and future large-scale research facilities. Covering computing architectures, software, data management, AI developments, and training, it aligns with FAIR data principles (Findable – Accessible – Interoperable – Reusable) and the UN's Sustainable Development Goals (SDGs).

While computing models across these fields vary significantly, a common challenge in the coming decade will be **scaling infrastructure to meet increasing demands**. Identifying synergies between research communities is crucial for achieving efficiency and sustainability. This includes international, transdisciplinary, diverse and open scientific collaboration as well as education and training in the field of computer science and the recognition of corresponding career paths. A **coherent international approach** is essential to address these challenges effectively.

Federated computing has evolved into a **powerful instrument for European integration and global scientific cooperation**, serving as a key enabler of science diplomacy. This white paper outlines key areas of action and recommendations to drive a unified, efficient, and future-proof computing strategy for European and international research infrastructures, with four key recommendations summarised below:

Computing Infrastructures

Sustained investment in computing and data infrastructures tailored to the specific scientific needs of these research communities remains essential. We recommend coordinated efforts to integrate High Performance Computing resources into research workflows, taking into account the strengths of current computing systems. This requires monitoring mechanisms to regularly assess the evolution of the computing needs and their impact on sustainable computational models and distributed computing infrastructure.

Software & Data Management

We recommend supporting software development and maintenance with significant investment to ensure performant and environmentally sustainable modern codes. Strong guidelines on Open Science and FAIR principles should drive community software roadmaps that identify critical tools and favour common solutions. Commonalities in distributed computing tools and services, as demonstrated by the EOSC Science Cluster ESCAPE, would help foster a unified FAIR data management ecosystem enabling economies of scale and promoting sustainability.

Artificial Intelligence

We recommend establishing a dedicated organizational structure to coordinate strategic investments in Artificial Intelligence for fundamental physics. We also recommend conducting a feasibility study to compare a centralized Graphics Processing Unit (GPU) facility with federated and hybrid High Performance Computing infrastructures. Additionally, we advocate for supporting physics-specific large language models, foundation models, and benchmarking while prioritizing scalable computing, robust data infrastructure, and Machine Learning operations.

Training & Careers

A vibrant training programme for collaborative computing requires continuous support. We recommend investing in people who can maintain and improve the soft- and hardware needed for our research to achieve our science goals while meeting Open Science and FAIR data management requirements. We strongly recommend implementing effective measures for training, as well as for recognizing achievements and promoting diversity across all relevant career paths. The broader community will benefit from open science approaches and a centralised collection of training materials and events. This needs to be supplemented by curricula and training pathways at universities and schools to acquire specific skills, including software and Artificial Intelligence.

1 Introduction

At the Joint ECFA-NuPECC-APPEC (JENA) Seminar in May 2022 in Madrid (JENAS 2022), both the plenary presentations and the closed session of funding agency representatives revealed that there is an increased need for discussions on the strategy and implementation of European federated computing at future large-scale research facilities. The computing landscape for large European infrastructures is diverse and lacks coherence. Particle physics explores adapting the WLCG model for HL-LHC, nuclear physics relies on facility-based computing with limited national center access, and astroparticle physics uses varied distributed models. Scaling challenges will dominate the next decade, making synergies crucial. To address this, APPEC, ECFA, and NuPECC launched a European initiative to explore federated computing strategies, aiming to identify resource and software synergies for future needs.

The JENA Computing Initiative (JENA Computing) first organized a workshop in (Bologna 2023). The conclusion of the workshop was that a more comprehensive assessment of computing needs for the next decade is necessary, also taking into account such important aspects as industrial participation, cybersecurity and future technologies such as quantum computing. This evaluation should be carried out in particular with regard to synergies that can benefit all three ENA communities as well as neighbouring research areas such as astrophysics or cosmology. The workshop was followed by thematic working groups, which led to specific Working Group Reports on the topics (i) federated infrastructures, (ii) software and heterogeneous architectures, (iii) data management and virtual research environments, (iv) machine learning and artificial intelligence, and (v) training, dissemination, education. The reports are summarised in this White Paper, supplemented by some some recommendations for more sustainable computing in the future. The JENA Computing Initiative as an integral part of the synergetic JENA activities will support the implementation of the recommendations given in this document in the long term and will continue to accompany the developments in the field of federated computing with regular evaluations.

2 HPC and HTC Infrastructures

This chapter focuses on identifying the needs and recommendations in the relationship of the High Throughput Computing (HTC) systems, used to manage and analyze the data of most of the large experiments such as the WLCG, with High Performance Computing (HPC) centres, and the integration among the two.

The analysis benefits from synergies with other initiatives, including a working group on HEP-HPC integration. A strategy meeting with EuroHPC (The European High Performance Computing Joint Undertaking) representatives took place in Ljubljana in September 2024. The meeting reported on the outcomes of a similar HEP-HPC strategy meeting held in the US earlier that year, which sought to leverage commonalities between the two regions. Additionally, a joint Europe-America-Asia strategy meeting has been organized at CERN in January 2025. Another key initiative is the SPECTRUM EU-funded project to design and distribute a survey to a “Community of Practice” closely aligned with the ENA community. SPECTRUM brings together selected players from our scientific fields, as well as from the area of current and future e-infrastructures (HPC, clouds, quantum computing). The outcome of the survey provided valuable insights into the current and future requirements of the ENA community. In all observations, care was taken to emphasise commonalities among the ENA communities, but also to take into account specific differences in needs.

Technical areas were identified where work will be needed to enable the effective integration of HPC centers and experimental facilities: edge services, federated access and AAI, workflow management, wide-area networking, data management, software deployment, programming models. Organisational areas were identified as well, in particular resource access and training and talent retention.

Distributed computing systems have become essential for large physics experiments, effectively handling the massive data volumes and enabling collaboration among geographically dispersed researchers. Analyzing this data often requires complex modeling and simulation, demanding immense computational power. At the same time, there is a growing trend toward centralizing computing resources in HPC centers. This centralization offers significant benefits, including improved energy efficiency and economies of scale, driving scientific communities to integrate HPC systems with their existing distributed data and workflow management infrastructures. By combining experiments, data facilities, theoretical physics needs, and HPC centers into a unified ecosystem, researchers can accelerate scientific discovery and strengthen international competitiveness.

However, several challenges remain that need to be addressed to fully realize this potential. The following recommendations outline key actions needed to achieve effective integration.

Recommendation: Sustained investment in ENA computing and data infrastructures tailored to the specific scientific requirements of these research communities remains essential.

Research activities within ENA are for a significant part driven by communities with moderate and diverse computing requirements. There is a significant risk that the priorities of EuroHPC centers may not align with the research-driven needs of these communities. Also, it should be recognised that the datasets produced in the ENA communities are a unique asset that will always need a dedicated infrastructure that cannot be outsourced. This infrastructure should also supply a constant baseline compute contribution.

Recommendation: A coordinated effort is needed to submit EuroHPC access requests from the ENA communities to develop a coherent roadmap for efficiently utilizing HPC systems.

This effort aims to achieve two key objectives: first, to establish a roadmap for addressing technical gaps, and second, to provide a consistent picture of global needs in preparation for a Strategic Access application. The ultimate goal is to be recognized as a strategic activity, enabling multi-year allocations of compute and storage at HPC facilities so that these resources can be incorporated into the long-term capacity planning of research groups and collaborations. At the same time, smaller-scale regional and national programs should continue to be exploited for the benefit of the ENA communities.

Recommendation: Securing adequate funding for software development in order to effectively integrate HPC resources into research workflows.

Substantial efforts are required to optimize scientific codes for new architectures, and to develop workflow and data orchestration software. Achieving long-term access to HPC resources could potentially free up funds, allowing them to be redirected toward these essential software development efforts.

Recommendation: Collaboration mechanisms, both at the technical and political level, should be identified in order to ensure that the future global scientific computing landscape builds on the strengths of current HPC and HTC systems.

With the unroll of initiatives like the Federation Platform and the AI Factories, EuroHPC JU systems are evolving towards a more interconnected and data-centric architecture. This evolution offers an opportunity to build on the successful experience of programs like WLCG in distributed data analysis systems. A key challenge in this transformation is the need for standardized, mutually agreed-upon policies across HPC centers to streamline access to these infrastructures. These policies should address critical areas such as authentication and authorization (AAI), data transfers, networking, edge services, and container support. Equally important is the establishment of processes that ensure the requirements of the communities are considered in the design and procurement of future HPC systems at regional, national, and European levels.

Recommendation: Establishing monitoring mechanisms to regularly assess the evolution of the CPU and GPU needs of the scientific collaborations and their impact on the computing models and on the distributed computing infrastructure.

There are large uncertainties in predicting the GPU needs in fundamental physics. It will be important to establish mechanisms to regularly assess the situation. These mechanisms should use agreed-upon metrics to track GPU usage and update resource estimations. The results should be incorporated into the roadmap for integrating HPC resources.

3 Software and Heterogeneous Architectures

The scientific communities are continuing to advance and are facing unprecedented software challenges due to growing data volumes, complex computing needs, and environmental considerations. As new experiments emerge, software and computing needs must be recognised and integrated early in design phases.

Across the scientific domains there are very significant challenges in software and computing, driven by ambitious physics programmes that deliver new detectors and observatories with increased data rates and data complexity. Software that is used directly to produce and process our science data, and to operate the corresponding infrastructure are considered in the following. The software to support these instruments, which is often very specific, is frequently ageing and needs investment, or replacement, which is particularly challenging for smaller experiments. The software is driven by a necessity to adapt to use modern hardware platforms, where high levels of parallelism are needed, particularly to execute efficiently on devices such as GPUs. The observational data, along with simulated data that models physics and detectors, needs to be reconstructed, analysed, increasingly in a distributed context, across sites and utilising facilities. Then software must be made available according to FAIR principles, which brings additional costs to genuinely achieve open science for these experiments. We reviewed the upcoming challenges, identify best practices, and recommend how to bolster support for widespread adoption of effective solutions to the needs.

From the JENA/SPECTRUM survey 90% of respondents reported that improving their software performance is an issue, due to data rates, physics goals and resource shortfall; additionally 58% reported lacking the necessary resources to do so. Here we identify specific positive policy decisions that, backed up by investment, will help. We make note in particular of the need to support access to important previous results in our fields, such as HEPdata or the proper management of nuclear physics data, which is a combined software and data preservation/access challenge. As software is so essential for our science domains, it is imperative to improve the recognition of software developers and support for their careers. We have identified a number of measures that can help to overcome the overarching challenges:

Recommendation: Develop strong software policies aligned with Open Science and FAIR Principles and provide resources to support those policies.

Only with appropriate software is the value of open data unlocked, so FAIR data and software principles should be expected and rewarded, e.g., by clear policies and support from laboratories and host institutes. This practice needs to be based on, and develop, appropriate standards with community and expert input.

Recommendation: Create Software Management Plans and road maps, to identify critical software, to be supported by adequate investment in software maintenance, refactoring, and development of modern solutions that help both reduce environmental impact and tackle new scientific challenges.

The maintenance of current software in the face of system upgrades and architectural diversity needs low level, but consistent, support. Yet, there is also a great need for radically improved software vital to manage the data complexity and rates from future facilities. Modern hardware demands the use of parallelism for efficiency, where multi-threading is usually also required in order to manage memory consumption. As platforms diversify, more effort is then needed to maintain build and validation systems. To make use of the heterogeneous resources expected at European computing sites, such an investment is vital.

Recommendation: Strategically invest in software that serves multiple experiments and disciplines and optimises data and workflow management.

Insofar as software can meet common goals for multiple users, well engineered systems deserve enhanced investment, particularly given the software engineering demands above. Unlocking the potential of distributed computing, especially for smaller endeavours, needs better, easier to use tools for data and workflow management.

Recommendation: Invest in training and reward trainers, adapting to new technologies and techniques as they arise.

The range of skills required to develop effective, scalable software solutions is very broad, with machine learning, data science techniques, high-performance computing and DevOps all playing a part (and new technologies will continue to emerge in the future). This requires quality training material and investment from program pilots, lessons authors and maintainers, instructors and helpers. Such efforts need to be properly recognised as valuable so that people will be motivated to develop, support and teach these skills.

Recommendation: Invest in software and computing work providing suitable career paths for software experts. This work must be seen as an integral part of our experiments and as first-class representatives of research. It is imperative that investment covers viable career paths for these staff, otherwise their commitment to our field will diminish.

To achieve our science goals in a way that also meets the requirements of Open Science and FAIR requires investment in people who can write, maintain and improve the software that is required by our collaborations. New roles are emerging that can be considered as viable for software experts - in particular the position of research software engineer.

4 Data Management and Virtual Research Environments

This chapter consolidates the strategic vision and actionable recommendations from the ESCAPE Science Cluster regarding federated data management, virtual research environments, and FAIR/open data principles. The ESCAPE collaboration vision outlines pathways to strengthen synergies among the ENA scientific communities. The focus is on evolving tools and services to meet the next decade's challenges, particularly in the areas of distributed data management, data access frameworks, and identity trust frameworks.

Federated Data Management: To develop unified guidelines and promote collaborative efforts for adopting Large Scale Distributed Data Management Systems built on shared underlying technologies. Designed to meet the evolving needs of Research Infrastructures (RIs): fostering interoperability, leveraging economies of scale and integrating sustainable practices to embrace FAIR Open Science Principles.
To address the growing demands of ESCAPE Research Infrastructures in terms of data volumes, number of files and sites (distributed computing), it is essential to adopt a Large-Scale Distributed Data Management System. The adoption of a common (or highly similar) system across sciences with well defined software, tools, and protocols would significantly enhance economies of scale and promote interoperability and sustainability.

Recommendation: Promote collaboration to adopt common Federated Data Management Systems capable to scale to the many-exabytes level and support entire data lifecycles.

Moreover, this approach would benefit both the software and computing teams of experiments and the infrastructure providers, including supporting laboratories, grid-like sites, cloud providers, and HPC centers. Such a system should integrate advanced metadata functionalities and seamlessly support the entire data lifecycle, from raw data ingestion to multi-tiered asynchronous data transfers. It should accommodate for global replication rules, enforce robust access policies, and interface efficiently with both metadata systems and diverse infrastructure providers. By leveraging metadata, the system allows efficient data discovery, data processing and fosters cross-infrastructures integration to leverage computing resources with RIs data repositories.

Recommendation: Ensure interoperability across computing providers, including HPC, by supporting the required protocols, adopting emerging standards and exploring edge services.

The proposed strategy builds upon the ESCAPE prototype Data Lake, fostering collaboration across participating RIs to enhance its capabilities. A key objective is to embrace data FAIRness by incorporating global replication rules, access policies (including data embargos and Open Data), enabling long-term Data Preservation, providing support for full data life-cycles, and ensuring flexibility for integration with diverse data processing infrastructure providers, including high-performance computing (HPC) centres and cloud services.

Virtual Research Environments, Data Access and Analysis Framework for Scientific Computing: To deliver a baseline framework of shared methodologies and best practices for data access and analysis to be used within ERIs.
The platform should integrate data discovery and analysis services – ranging from interactive computational notebooks and visualization tools to bulk data processing systems – with large-scale data management services, such as Data Lakes and content delivery networks. The proposed Data Access and Analysis Framework aims to empower European researchers within

RIs by providing a unified model for data access and analysis, supporting a collaborative and efficient research environment while addressing the critical aspects of reproducibility and long-term analysis preservation.

> Recommendation: Support the development of a modular, scalable analysis platform that RIs can build on and customize, ensuring software sustainability and analysis reproducibility.

The RIs identified the interest to enhance their interactive analysis capabilities, lowering the barrier entry for newcomers and do first analysis, promoting visual data integration with live algorithms (live histogram generation on the notebook), and empowering end-users under the scope of Open Data, Citizen Science and Learning activities. The system should ensure seamless access to experiment's data and software, and fosters reproducibility and analysis preservation through an integrated, user-friendly platform.

Federated identities and trust framework: The objective is to grant seamless and secure access to data management services, analysis facilities, and resources, thus fostering collaboration and interoperability. The framework aims to enhance the overall cybersecurity posture of RIs, protect sensitive information, and facilitate cross-institutional collaboration and threat intelligence. The activities should be scoped in alignment with related Authentication and Authorization Infrastructure (AAI) federation activities in the EOSC and other Science Clusters to provide a harmonized integration of trust frameworks based on common standards.

> Recommendation: Promotion of a common model of identity trust across Research Infrastructures.

A Federated Identities and Trust Framework is essential for standardizing practices around authentication, authorization, and identity management within the European research community. By creating a common layer of trust, RIs can streamline access to resources, foster collaboration, and ensure the security and integrity of shared identities across a diverse array of platforms. The recommendation is aligned with ongoing initiatives under EOSC and the European Science Clusters, leveraging their efforts in Authentication and Authorization Infrastructure (AAI) federations.

A set of more detailed recommendations has been identified as priority topics from the ESCAPE community, to be used as building blocks to define a roadmap for the coming years and are described in the long document.

5 Machine Learning and Artificial Intelligence

Artificial intelligence (AI) is transforming scientific research, with deep learning methods playing a central role in data analysis, simulations, and signal detection. AI integration is advancing steadily, however, broader adoption remains constrained by challenges such as limited computational resources, a lack of expertise, and difficulties in transitioning from R&D to production. Here we provide a summary of a strategic roadmap,informed by a community survey, to address these barriers. It outlines critical infrastructure requirements, prioritizes training initiatives, and proposes funding strategies to scale AI capabilities across fundamental physics over the next five years.

The ENA community's ability to conduct impactful AI research depends on scalable high-performance computing resources. Traditional CPU-based infrastructures in high-energy physics are no longer sufficient, necessitating a shift to GPU-based systems optimized for parallel computing and AI workloads. Two main strategies could address this need: establishing a centralized, large-scale GPU facility or expanding existing HPC infrastructure with additional GPUs and cloud integration. Initiatives like the EuroHPC AI Factories program may help address funding and implementation challenges.

> Recommendations: Convene dedicated discussions with national research groups and funding bodies to assess and compare the feasibility of a centralized large-scale GPU facility versus federated and hybrid HPC infrastructures. Establish a scalable data infrastructure initiative by creating shared repositories and tools, and developing platforms for distributed workloads.

Despite significant progress, AI research in physics remains fragmented, with most efforts focused on proof-of-concept applications rather than fully integrated, production-ready solutions. However, these advancements are often driven by a small number of researchers, as scaling AI solutions beyond proof of concept remains challenging.

> Recommendations: Encourage funding to transition AI-driven R&D activities into production-ready applications within established experimental workflows. Allocate dedicated funding to establish and support specialized Machine Learning Operations (MLOps) personnel.

With many researchers already using tools like ChatGPT, it is important to consider both the benefits and limitations of these commercial AI models in the context of fundamental physics. Establishing dedicated funding schemes and a collaborative structure to develop community-driven large language models and foundation models trained on domain-specific data would allow for customization, improving transparency and control. These models could learn meaningful representations, serving a large variety of downstream tasks. To succeed, we need to invest in collaborative initiatives to share data, resources, and GPU capacity. Additionally, a dedicated effort should be made to develop and maintain extensible benchmarks for various AI tasks in fundamental physics, such as event classification, parameter inference, tracking, and anomaly detection.

> Recommendations: Invest in the creation of "science Large Language models (LLMs)" tailored to the unique challenges of fundamental physics and science. Develop community-driven foundation models trained on domain-specific data to learn meaningful representations serving a large variety of downstream tasks. Develop and maintain extensible benchmarks for various AI tasks in fundamental physics.

The AI community is increasingly aware of the environmental impact of large-scale model training, highlighting the need for improvements in both algorithms and hardware to reduce energy consumption and training time. Ongoing research in "green AI" is exploring ways to optimize performance metrics and reduce computing costs, including selecting low-carbon-emission hardware and optimal compute locations. Full lifecycle assessments for hardware, transparency regarding power and hardware resources, and novel neuromorphic schemes for power-constrained environments could further reduce emissions. Furthermore, integrating FAIR principles into AI research promotes sustainability by reducing duplication of research efforts and improving resource efficiency.

Recommendations: Investigate and adopt benchmarks that are suitable for fundamental sciences to raise awareness of the environmental impact of large AI models. Cooperate with infrastructure and computing sites to minimise carbon costs of compute-intensive AI tasks. Develop activities aiming to integrate FAIR compliance into publication criteria and practices.

To address the challenges of AI in fundamental physics, a coordinated structure, inspired by initiatives like the European Coalition of AI for Fundamental Physics (EuCAIF), shall be established. This would streamline resource allocation, encourage interdisciplinary collaboration, and accelerate AI adoption. Supporting this, funding for training courses and summer schools will equip researchers with the skills to implement open research and reproducibility practices, while industry partnerships offer placements and event sponsorships. Interdisciplinary initiatives should also unite physicists, AI specialists, and HPC experts, promoting cross-domain knowledge transfer through workshops and open-source collaboration.

Recommendations: Establish and support a dedicated organisational structure to coordinate strategic investments in AI for fundamental physics. Fund the development and organization of practical training courses and summer schools. Establish interdisciplinary research initiatives that bring together physicists, AI specialists, software engineers and HPC experts. Provide dedicated funding to support cross-domain knowledge transfer.

6 Training, Dissemination, Education

This chapter examines the current status and measures for education and training for Federated Computing and developed recommendations on how to proceed in this direction.

The **training** of emerging scientists and software specialists is a vital mission for the European communities supported by JENA. As research infrastructures, experimental setups, theories, and transnational collaborations become more complex and expansive, there is an increasing need for targeted training. At the same time, since the stall of computer clock frequencies in the mid 2000s the computing landscape has undergone major changes with e.g. the advent of multi-core CPU processors and specialised hardware accelerators such as GPUs. A large fraction of future computing resources are likely to be provided by HPC which receive most of the computing power through hardware accelerators such as GPUs or FPGAs, which the community needs to adapt to. The use of external resources like clouds or HPC as well as the managing and processing of very large data volumes (Exabytes scales) requires dedicated user training activities. This specialized approach includes familiarizing trainees with essential methods, instruments, software tools and the operation of significant computing facilities, as well as data acquisition systems and analysis protocols.

Recommendation: A vibrant program of trainings on Federated Computing to satisfy the needs of the ENA communities requires continuous support in the future. The community will strongly benefit from a centralised repository of training materials and events, enhanced with proposed curricula and training pathways to achieve certain proficiency.

Diversity in the IT workforce is an important aspect that needs to be valued and continuously supported, alongside equity and inclusion. A key impediment in the **career and capacity building** of Federated Computing talent is the existing divide between scientific and computing support staff in most working environments, from national laboratories to universities.

The standard scientific career path is well defined and often relies on excellence at the scientific level, in terms of quality and quantity of publications, scientific funding and other parameters that may directly or inadvertently exclude service to Federated Computing infrastructures; software design, development or documentation. Scientists with strong software experience may have difficulties in developing competitive curricula compared to those that simply rely on software tools. In parallel, we identify a second group of relevant staff, which we name here "research software engineers". They typically belong to computing support groups (as opposed to scientific research groups), and their work is usually focused around software design and documentation, or hardware exploitation. Such scientific software engineers are invaluable for new scientific developments, but the corresponding career paths are often not as well defined. Industry can easily be a more attractive destination in view of this professional instability and generally lower competitive salaries.

Recommendation: JENA should support the creation or the exploitation of existing forums like the HEP Software Foundation for discussions and developments such as conferences, workshops, seminars for computing aspects open to all communities in order to retain current and attract new talent.

Recommendation:
The competitive economic environment in the digital sector, together with a lack of stability in some career paths, reduces the capacity to attract or keep the best talent in Federated Computing. It is important to promote and support competitive salaries and attractive career opportunities.

Dissemination beyond the ENA scientific communities, towards society, industry and other stakeholders is also a key communication activity. This is particularly important at a time where science, the scientific method and the scientific community are under intense scrutiny by the general public on several relevant fronts. While Federated Computing alone may not be easily amenable to outreach, there are several activities that have been enabled by computing infrastructure.

Recommendation: The promotion of open science is a priority for the ENA community and should be fully incorporated into training programmes and activities.

The focus here is on Higher Education at University level, as we assume that this is the minimum level required for Federated Computing awareness, understanding and practice. In spite of the importance of Federated Computing, HPC and other advanced computer science tools (such as Machine Learning) for scientific developments, there is a conspicuous lack of University-level provision of courses on these tools, particularly in scientific degrees.

Recommendation: With university students in particle, astroparticle and nuclear physics not necessarily being exposed to software and computing techniques which are required in the respective ENA areas it will be beneficial if the required skills education can be passed to the universities and schools. Among many relevant skill sets the feedback on education in Artificial Intelligence and Machine Learning is of utmost importance. This can be done either directly with universities and engineering schools or indirectly on a political level through the national sciences ministries.

Recommendation: Attending training courses in many cases is mandatory for students e.g. while attending a PhD program. With many training and schools being offered within the ENA communities the granting of ECTS points for attending those will be beneficial for the students. A generalised approach of granting those points for attendance will be helpful to the training communities.

7 Sustainability in Computing

The central goal of sustainability is to reconcile economic development, social justice and the conservation of natural resources. This applies both to research in general and to computing as a service for basic research (in the ENA communities). We understand computing here in a broader sense, which includes all nodes of the data lifecycle of scientific data, as each step/node requires sophisticated computing. This includes in particular the handling of FAIR data, as the sustainability of the data lifecycle can only be discussed on the basis of FAIR data. All the measures described for maintaining or improving the data-intensive computing capacities in particular with regard to sustainability require corresponding funding and should be supported by appropriate funding programs. The deliberations and discussions led to the following recommendations for initial strategic measures that should not only strengthen the sustainability of research, but also contribute to sustainability research:

Recommendation: The cultural shift towards sustainable computing should be actively supported, e.g. by monitoring key performance indicators to be determined using a guideline, evaluating structural boundary conditions and through targeted promotion of knowledge transfer, as well as education and training in sustainability.

Recommendation: The energy and resource efficiency of IT infrastructures should be presented transparently and must continue to be increased. This applies equally to the design, energy supply, operation and use of infrastructures. Activities with regard to increasing efficiency and the associated development of eco-systems for monitoring, accounting and usage orchestration, including prediction of required energy consumption and available energy mix, needs to be carried out.

Recommendation: The lifecycle analysis of the required digital devices, the total costs - acquisition, maintenance, repair costs, environmental impact, workload - must be optimized, which requires the development of an action plan for implementation.

Recommendation: Huge data streams can only be managed if the fraction of information that is actually relevant is extracted during data collection. The R&D of corresponding algorithms and processes will make substantial contributions to effective, FAIR and therefore sustainable data analysis and data reduction and should be promoted.

Recommendation: Software development, which plays a central role in the sustainable use of resources, should be promoted more intensively, also in order to train young scientists with the necessary dual expertise in methodological and domain knowledge.

Glossary

AAI: Authentication and Authorization Infrastructure
AI: Artificial Intelligence
AI Factories: EU project to leverage EuroHPC to develop cutting-edge AI models
APPEC: AstroParticle Physics European Consortium
ARM: Advanced RISC Machine
CERN: Centre Européen de Recherche Nucléaire
CHEP: Computing in High Energy and Nuclear Physics (conference)
CPU: Central Processing Unit
DevOps: Term for combining software development with testing and IT operations
ECFA: European Committee for Future Accelerators
ECTS: European credit transfer and accumulation System
ENA: ECFA-NuPECC-APPEC
ENVRI-FAIR: ENVironment Research Infrastructure-FAIR
EOSC: European Open Science Cloud
ESCAPE: European Science Cluster of Astronomy & Particle Physics ESFRI RI
ESFRI: European Strategy Forum on Research Infrastructures
EuCAIF: European Coalition of AI for Fundamental Physics
EuroHPC: short form of EuroHPC JU
EuroHPC JU: The European High Performance Computing Joint Undertaking
FAIR: Findability, Accessibility, Interoperability, and Reusability
FPGA: Field-Programmable Gate Array
GPU: Graphics Processing Unit
HEP: High Energy Physics
HEPdata: Repository for publication-related High-Energy Physics data
HL-LHC: High Luminosity - Large Hadron Collider
HPC: High Performance Computing
HSF: HEP Software Foundation
HTC: High Throughput Computing
IT: Information Technology
JENA: Joint ECFA-NuPECC-APPEC
LHC: Large Hadron Collider
LLMs: Large Language Models
ML: Machine Learning
MLOps: Machine Learning Operations
NuPECC: Nuclear Physics European Collaboration Committee
OSG: Open Science Grid
PANOSC: Photon And Neutron Open Science Cloud
R&D: Research and Development
RI: Research Infrastructure
RISC: Reduced Instruction Set Computer
SDG: Sustainable Development Goals
SPECTRUM: EU project to elevate data-intensive science in Europe
SSHOC: Social Sciences & Humanities Open Cloud
WLCG: Worldwide LHC Computing Grid