SPECTRUM

CONNECTING THE DOTS
ISC High Performance
JUNE 10 – 13, 2025 | HAMBURG, GERMANY

BOF: Towards a Strategy for Future Research Infrastructures

www.menti.com
Code:

Sergio Andreozzi
EGI Foundation (NL)
SPECTRUM Project Director

Hans-Christian Hoppe
Senior Project Lead, FZJ/PARTEC AG (DE)
SPECTRUM Project WP5 Leader

Maria Girone
Head of Openlab, CERN
SPECTRUM Project WP6 Leader

Christian Engelmann
Distinguished Scientist and Group Leader,
Intelligent Systems and Facilities
Oak Ridge National Laboratory (USA)

Sandra Diaz Pier
Scientific lead of the Simulation and Data Lab Neuroscience, FZJ (DE)

Jeff Wagg
Radio Astronomy Data Scientist, OCA (FR)
SPECTRUM Project WP4 Co-Leader

Funded by the European Union

SPECTRUM is funded by the European Union
Grant Agreement Number 101131550

BOF Objective

CHALLENGE

Next-generation scientific instruments (e.g. High-Luminosity LHC, Square Kilometer Array, LOFAR 2.0) will generate data volumes requiring compute capabilities orders of magnitude beyond current resources

SPECTRUM PROJECT

Developing a strategic research and innovation agenda and Technical Blueprint for Compute and Data Continuum for High-Energy Physics and Radio Astronomy

GOAL OF THIS SESSION

Engage the ISC Community and gather feedback to enrich the work

BOF Structure

Introduction and Statements

- **Introduction to the session**
  - Sergio Andreozzi, EGI
- **SPECTRUM Project Overview**
  - Sergio Andreozzi, EGI
- **Vision from Radio Astronomy**
  - Jeff Wagg, OCA
- **Vision from High-Energy Physics**
  - Maria Girone, CERN
- **Vision from NeuroScience**
  - **Sandra Diaz, EBRAINS**
- **Vision from Interconnected Science Ecosystem (INTERSECT) / Integrated Research Infrastructure (IRI)**
  - Christian Engelmann, ORNL, US

Panel + Audience

- **Moderator: Hans-Christian Hoppe**
- **Scientific Domain**
  - **Jeff Wagg, Radio-Astronomy**
  - **Maria Girone, HEP**
  - **Sandra Diaz, NeuroScience**
- **e-Infrastructures**
  - **Christian Engelmann, IRI**
  - **Sergio Andreozzi, EGI**
- **Audience engagement**
  - Via Mentimeter

How to Participate during the BOF

Go to

www.menti.com

Enter the code

7770 7562

Or use QR code

SPECTRUM Project Overview

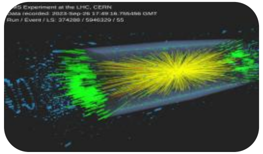
Sergio Andreozzi, EGI Foundation
SPECTRUM Project Director

SPECTRUM

Elevating Data-Intensive Science in Europe

CHALLENGE

The **computing** for the **next generation of scientific instruments** in High-Energy Physics and Radio Astronomy is an **unsolved problem**

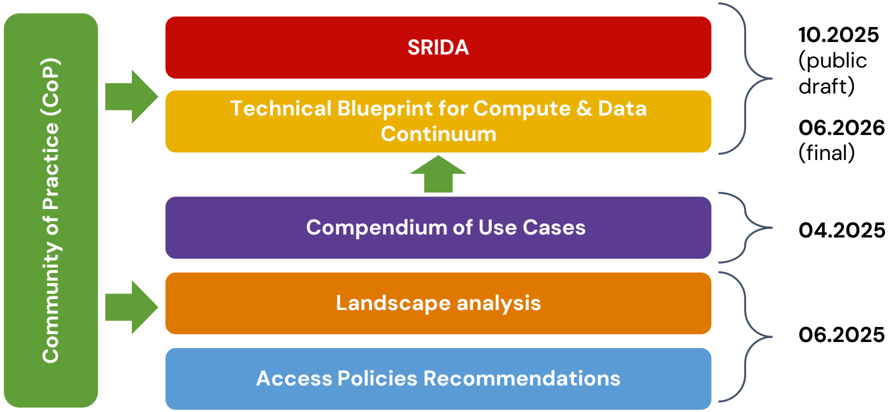


GOAL

Deliver a **Strategic Research, Innovation and Deployment Agenda (SRIDA)** and a **Technical Blueprint** for a **European compute and data continuum**.

IMPACT

- **Common vision** on future of infrastructures for data-intensive research
- **Aligning efforts** and reduce fragmentation among initiatives
- **Optimising use** of common digital infrastructures for research



Research Communities

CERN, INFN, ASTRON, CNRS, Observatoire de la Côte d'Azur, Neovia Innovation

e-Infrastructures

EGI, Jülich Forschungszentrum, CINECA, SURF

The SPECTRUM Project is funded by the European Union Grant Agreement Number 101131550 - **More Info: spectrumproject.eu**

Current Results

https://www.spectrumproject.eu/publications

Community of Practice (CoP)

- Cross-disciplines WGs with experts from HEP, RA and digital infras (link)
- Knowledge Base (link)
- Collaboration platform


SPECTRUM

D3.1 Community of Practice – Interim report

Status: UNDER EC REVIEW

Dissemination Level: Public

Current Working Groups

- WG1: Data Management and Access
- WG2: Workflow management and organization
- WG3: Compute Environment
- WG4: SW tools
- WG5: Scientific Use cases
- WG6: Facilities

Compendium of Science Use Cases (15)

- From science case to technical challenges, requirements, gaps
- Both technical and policy aspects


SPECTRUM

D5.1 Representative use cases: analysis and alignment

Status: UNDER EC REVIEW

Dissemination Level: Public

- High Energy Physics (7)
- Radio Astronomy (5)
- Theoretical Physics (1)
- Chemistry (1)
- Meteorological Science (1)
- Neuroscience (1)

Upcoming Results

https://www.spectrumproject.eu/publications (**to be published next week**)

Access Policies Recommendations	Landscape Analysis
• Existing access policies across the continuum • Development of recommendation for Interoperable access	• Existing approaches, services, technical solutions and policies for the federation of data and compute infrastructures


SPECTRUM

D5.2 Interoperable access policies: analysis and recommendations

Status: DRAFT

Dissemination Level: Public


SPECTRUM

D5.3 Landscape of RIs: technologies, services, gaps

Status: DRAFT

Dissemination Level: Public

Upcoming Results

Public Draft for Consultation: Nov 2025 – Final Version: Jun 2026

Technical Blueprint for Compute & Data Continuum

- Proposal for a compute-and-data continuum infrastructure design and architecture

SPECTRUM

D6.1 Technical Blueprint for Compute and Data Continuum

Strategic Research, Innovation and Deployment Agenda

- Vision and overall goals
- Research, innovation and deployment agenda
- Focus on data-intensive science and infras in HEP & RA

SPECTRUM

D7.1 SPECTRUM Strategy Research Innovation and Deployment Agenda

Vision from Radio Astronomy

Jeff Wagg
Radio Astronomy
Data Scientist, OCA
SPECTRUM Project WP4 Co-Leader

Vision from Radio Astronomy (RA) domain

Jeff Wagg (CNRS/OCA)

Preparing for Exabyte scale computing

- Radio astronomy is not new to the problem of big data processing, but the bar is rising rapidly (in the early days of the Very Large Array, data were written to tape and correlated away from the site with typical file sizes <100 MB)
- Big change in recent years has been the transition of researchers processing their data multiple times on a single laptop, to having to rely on HPC resources and pipelines run once, or twice - need to socialise the community, who will also have to accept some pre-processing of their data by observatory staff
- A related paradigm shift is in the disposal of raw data, which radio observatories have traditionally archived indefinitely. This will no longer be possible for SKAO, which will produce ~200 PB/year of science data products
- Wide-field radio astronomy datasets (1000s of square degrees with much greater sensitivity), including both continuum and spectral line, will require automated solutions for discovery and source characterization

Future Solutions

- Radio astronomy is following the path of HEP in adopting federated computing solutions with data storage and compute distributed across multiple heterogeneous (international) nodes
- Users will process the data through a common science gateway that is independent of physical location and use edge processing to analyse their data
- Users will adopt regional center network to generate advanced data products, including AI tools for tasks such as source extraction and characterization.

Vision from High-Energy Physics

Maria Girone
Head of Openlab, CERN
SPECTRUM Project WP6 Leader

HEP domain vision: Towards an exabyte-scale compute & data continuum

Maria Girone (CERN)

Exponential data growth

- Next-generation HEP instruments (e.g., HL-LHC upgrades) will generate exabytes of data over the coming decade, requiring seamless federation of storage and compute resources across Europe.
- HEP workflows span from large-scale batch reconstruction to interactive analysis; a unified continuum ensures HPC/HTC is leveraged efficiently.

Integrated HPC & Cloud utilization

- Standardized data access, network, and AAI across HEP and HPC sites are crucial: current site-specific policies must evolve toward interoperable, standardized interfaces.
- Cloud resources complement HPC by providing elasticity for bursty workloads; however, long-term cost models (especially egress charges) necessitate optimized data placement strategies.

Federated, domain-aware infrastructure

- Build upon the WLCG experience to federate heterogeneous resources (HTC, HPC, Cloud) under a common workload-management layer, enabling dynamic scheduling based on data locality and compute capability.
- Establish a European compute-and-data continuum that transparently exposes resources to HEP experiments, minimizing per-site bespoke integration efforts and reducing operational overhead.

Key HEP priorities & strategic directions in SPECTRUM

Maria Girone (CERN)

Portable, heterogeneous software

- The era of “one-architecture dominance” has ended: Support x86_64, ARM, RISC-V, GPUs via unified frameworks (e.g., Kokkos, Alpaka).
- Consolidate on a few community-backed programming models to avoid repeated porting.

AI/ML integration for accelerated discovery

- Complement traditional CPU-intensive simulation and reconstruction workflows with AI/ML-based and data-driven approaches.
- Build AI-optimized environments (GPUs, accelerators) and train HEP experts to deploy models.

Sustainable ecosystem & careers

- Advocate for long-term funding of operations to lower integration overhead.
- Cultivate rewarding career paths for computing experts, critical for multi-decade experiments, and recognize computing contributions to attract and retain talent.

Vision from NeuroScience

Sandra Diaz Pier
Scientific lead of the Simulation and Data Lab Neuroscience, FZJ

Vision from neuroscience (EBRAINS) domain

Sandra Diaz (FZJ)

Heterogeneity in implementation and maturity

- The different levels of maturity of the codes currently available in the neuroscience community make it difficult to have a one-fits-all solution for HPC support and further development.
- Moving from CPU to GPU requires considerable refactoring efforts in particular for some of the legacy community codes.

AI and neuroscience

- AI plays a critical role in our field two ways: AI for neuroscience and neuroscience for AI—this represents a great opportunity that we would like to exploit further and extensively using HPC.
- Neuroscience is a driver for the development of new accelerators e.g. neuromorphic computing, as well as algorithms for energy efficient AI.
- At the same time, foundation models trained on structural and functional data of the brain will definitely complement and change the way neuroscience is done in the next years – this will require substantial computational resources close to the data.

Vision from neuroscience (EBRAINS) domain

Sandra Diaz (FZJ)

Challenges on HPC

- Managing large, very heterogeneous data sets of experimental and synthetic data is complex and securing long term storage for these data sets is still a challenge.
- Memory management and communication play a key role in neuroscience applications, in particular with simulations of brain dynamics.
- Hybrid strategies including cloud and HPC are desirable to support new workflows and interactive access.

Community

- Our community is highly interdisciplinary—including experimentalists, clinicians, computational neuroscientists, electronic engineers, philosophers, etc.
- Networking, community building, training and education initiatives are essential to enable the neuroscience community to leverage the current and future computational resources.
- Making HPC easier to access and use with pre – configured software installations, containerization and workflows helps uptake by the community.

US View: Integrated Research Infrastructure & Federated Ecosystems

Christian Engelmann
Distinguished Scientist and Group Leader,
Intelligent Systems and Facilities
Oak Ridge National Laboratory (USA)

US View: US Department of Energy Integrated Research Infrastructure (IRI)

Christian Engelmann (ORNL)

- Will seamlessly link experimental and observational scientific user facilities, data assets, and advanced computing resources so that researchers can combine these tools in novel ways that radically accelerate discovery: https://iri.science
- 2022 IRI Architecture Blueprint Activity involving 150 experts from 28 DOE user facilities across 13 US national laboratories identified:
  - *6 practice areas:* User experience; resource co-operations; cybersecurity and federated access; workflows, interfaces, and automation; scientific data life cycle; and portable/scalable solutions
  - *3 patterns:* Time sensitive; data integration-intensive; and long-term campaign
- Ongoing effort involves primarily ALCF, OLCF, NERSC, and ESnet and will include HPDF in the future.

US View: ORNL Interconnected Science Ecosystem (INTERSECT)

Christian Engelmann (ORNL)

- A federated ecosystem that connects scientific instruments and robot-controlled laboratories with computing and data resources at the edge, the Cloud or the high-performance computing center to enable autonomous experiments, self-driving laboratories, smart manufacturing, and artificial intelligence driven design, discovery and evaluation: https://www.ornl.gov/intersect
- Build on science use case design patterns, a system of systems architecture, and a microservice architecture: https://intersect-architecture.readthedocs.io
- A Python-based software development kit: https://github.com/INTERSECT-SDK
- ORNL internally-funded R&D effort targeting several autonomous laboratories/experiments:
  - Robotic chemistry lab, additive manufacturing, electron microscopy, etc.

Discussion: Panel + Audience

Moderator
Hans-Christian Hoppe
Senior Project Lead, FZJ/PARTEC AG
SPECTRUM Project WP5 Leader

Mentimeter
Xavier Salazar
EGI Foundation (NL)
SPECTRUM Project
Exploitation Manager

Sergio Andreozzi
EGI Foundation (NL)
SPECTRUM Project
Director

Maria Girone
Head of Openlab, CERN
SPECTRUM Project WP6 Leader

Christian Engelmann
Distinguished Scientist and
Group Leader, Intelligent
Systems and Facilities
Oak Ridge National Laboratory

Sandra Diaz Pier
Scientific lead of the
Simulation and Data Lab
Neuroscience, FZJ

Jeff Wagg
Radio Astronomy
Data Scientist, OCA
SPECTRUM Project WP4
Co-Leader

Funded by
the European Union

SPECTRUM is funded by the European Union - Grant Agreement Number 101131550

www.menti.com
code: 7770 7562

How to Participate during the BOF

Go to

www.menti.com

Enter the code

7770 7562

Or use QR code

Panel Discussion

- What are the main compute & data challenges and gaps for your research community?

- How current technologies and policies should evolve to meet the identified challenges/gaps?

- What are the main priorities to tackle over the next 5-7 years?

www.menti.com
Code: 7770 7562

How can you stay engaged?

Email: **info@spectrumproject.eu**

Join the Community of Practice

https://www.spectrumproject.eu/spectrumcop

Follow us on LinkedIn

https://www.linkedin.com/company/spectrum-project-eu

Participate in the Public Consultation

Nov 2025 – Jan 2026

Subscribe to the Newsletter

https://www.subscribepage.com/spectrumproject

ENJOYED THIS BOF

Bookmark
it in your
personal
schedule to
rate it later.

Thank you!