SPECTRUM

# Strategic Research Innovation and Deployment Agenda (SRIDA)

May 2026

## Document Description

**Strategic Research Innovation and Deployment Agenda (SRIDA) – pre-final version**

| | |
|---|---|
| **Copyright Status** | © This material by Contributing Parties of the SPECTRUM Consortium is licensed under a Creative Commons Attribution 4.0 International License. |
| **Funding** | Funded by the European Union<br>SPECTRUM is funded by the European Union – Grant Agreement Number 101131550 |
| **Photo Credit** | The images are adjusted from pictures provided by CERN; OCA/CNRS; SKAO/Max Alexander |
| **Lead Partner** | EGI Foundation |
| **Project Website** | www.spectrumproject.eu |
| **Document Webpage** | https://www.spectrumproject.eu/srida |
| **Digital Object Identifier** | 10.5281/zenodo.20182879 |
| **Author(s)** | **Lead Editors:**<br>Sergio Andreozzi (EGI Foundation), Xavier Salazar (EGI Foundation)<br>**Contributors:**<br>Concezio Bozzi (INFN), Florian Berberich (FZJ), Luis Cifuentes (FZJ), Stefano Bagnasco (INFN), Luca dell'Agnello (INFN), Tommaso Boccali (INFN), John Swinbank (ASTRON), Chiara Ferrari (CNRS/OCA), Jeff Wagg (CNRS/OCA), Shan Mignot (CNRS/OCA), Raymond Oonk (SURF), Maria Girone (CERN), Eric Wulff (CERN) |
| **Reviewers** | Luciano Gaido (INFN), Mylène Pischella (NEOVIA |
| **Moderated by** | Patricia Ruiz (EGI Foundation) |
| **Design** | Eric Graventein (EGI Foundation), Gwen Franck (EGI Foundation),<br>Anabelle Tumboimbela (EGI Foundation) |

# Table of Contents

# Executive Summary

European data-intensive science is approaching a step change in scale. The instruments coming online over the next decade, including the High Luminosity Large Hadron Collider (HL-LHC), the SKA Observatory and LOFAR 2.0, will generate data volumes and processing demands an order of magnitude beyond current infrastructure capacity. The HL-LHC alone is projected to require tens of exabyte storage and tens of millions of CPU-equivalent cores from 2030 onwards; SKA Phase 1 will produce hundreds of petabytes per year before scaling further. These facilities operate on multi-decade timelines, so the architectural and governance choices made in the next five years will determine whether Europe can extract the scientific return on infrastructure investments already committed. Meeting this challenge requires coordinated action across scientific communities, computing and data service providers, and policy makers. **This Strategic Research, Innovation and Deployment Agenda (SRIDA) presents thirteen priorities, five investment areas, and a phased multi-annual roadmap that turn that coordination into action.**

Europe's research infrastructure landscape combines EU-level coordination, through ESFRI for long-term research infrastructure planning, EuroHPC for supercomputing capability, and EOSC for federated data and open science, with national investments that still provide the majority of operational storage and computing capacity. This agenda builds on operational experience from federated infrastructures already coordinating nationally-funded resources at European scale: EGI for national and research institution computing facilities, WLCG for LHC-related computing/storage facilities, the SKA Regional Centre Network (SRCNet) and GEANT for the networking. These existing federations show what is feasible and where the next generation of coordination must reach further.

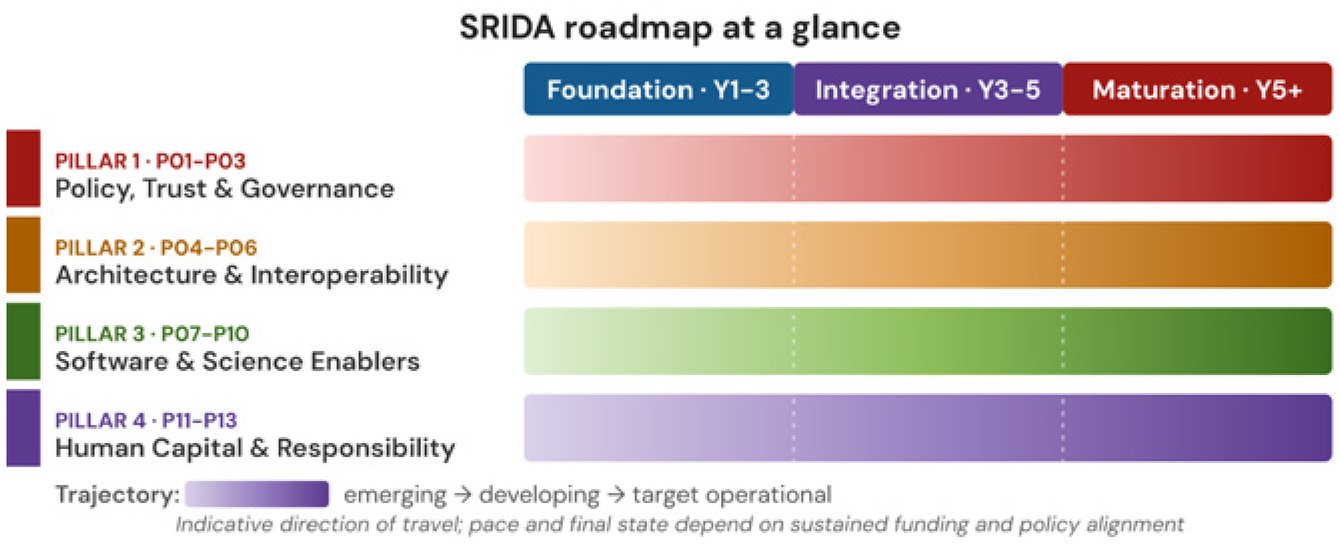
Seven drivers shape the strategic environment. Infrastructure Scale: flagship instruments will move from petabyte to exabyte operations, requiring architectural evolution across compute, storage, and networking, with heterogeneous accelerators (GPU, FPGA, DPU, and emerging quantum devices) becoming the default rather than the exception. AI Adoption: artificial intelligence is being adopted rapidly across the research data lifecycle, with most applications still moving from proof-of-concept to production and demand for AI-capable infrastructure growing across European research communities. Environmental Sustainability: energy availability, water use, and embodied carbon in hardware are becoming primary constraints on infrastructure growth, making efficiency a design parameter rather than an afterthought. Security and Trust: post-quantum cryptographic transitions, evolving authentication requirements, and stricter data governance frameworks add complexity to federated operations. Long-term Preservation: instrument timelines of 20 to 50 years require software, data, and AI models to remain usable across hardware generations, with FAIR principles extending to workflows and provenance. Workforce Capacity: skills profiles are shifting toward research software engineering, data engineering, and applied AI; career structures, recognition, and competition with industry for technical talent remain unresolved. Digital Sovereignty: European policy increasingly treats digital capacity as a strategic asset, requiring open technologies, data processing within European jurisdictions, and resilience against geopolitical disruption.

**Responding to these drivers requires progress against five strategic goals: coherent governance connecting thematic research infrastructures with horizontal e-infrastructures; seamless resource access through multi-year allocation and federated authentication; technical interoperability across heterogeneous computing and cross-facility workflows; sustainability across software portability and environmental impact; and human capacity built on expertise combining domain science with computing skills.**

The agenda organises thirteen priorities within four pillars to deliver these goals. Pillar 1: Policy, Trust and Governance establishes cross-infrastructure governance, multi-year resource allocation, and federated identity as foundations for everything that follows. Pillar 2: Architecture and Interoperability enables heterogeneous computing, federated data management, and cross-facility workflow orchestration. Pillar 3: Software and Science Enablers advances AI/ML in production, code portability and performance, scientific reproducibility, and the long-term preservation of data, software, and workflows. Pillar 4: Human Capital and Responsibility addresses community collaboration and co-design, environmental sustainability, and workforce development. The pillars are interdependent: technical interoperability without governance fails to scale, governance without skilled people cannot be implemented, and none of it is sustainable without explicit attention to environmental footprint.

Investment spans five areas: computing infrastructures, data infrastructures, software and tools, governance and coordination, and human capital. Open science and environmental sustainability cut across all five. Three categories are routinely underestimated and are essential for realising value from capital investments: coordination and federation costs, sustained software development beyond initial deployment, and long-term preservation of data, software, and workflows under FAIR principles. Effective investment depends on coherent action across EuroHPC, EOSC, Horizon Europe, the Digital Europe Programme, national research and infrastructure programmes, and operational funding for research infrastructures themselves.

Implementation follows three phases. The short-term foundation phase (1 to 3 years) establishes governance mechanisms, federated identity, and standardised interfaces, and begins co-design between thematic research infrastructures and e-infrastructures with the aim of aligning services and interfaces, not merging organisations. The medium-term integration phase (3 to




5 years) deploys heterogeneous computing capabilities at scale, federated data management, and workflow orchestration spanning HPC, HTC, cloud, and edge resources. The long-term maturation phase (5 years and beyond) achieves AI/ML at production scale, end-to-end reproducibility, and a stable career framework for research computing. Governance and identity foundations must precede technical integration; workforce development and environmental management run across all phases to keep the effort sustainable.

**The agenda places concrete tasks on each stakeholder group.** Scientific communities should publish multi-year resource plans, take part in co-design with e-infrastructures, identify high-impact use cases for moving AI/ML to production, and contribute to shared software and data stewardship. Service providers should establish formal liaison mechanisms with thematic research infrastructures, deploy federated identity and standardised interfaces, expose heterogeneous resources through common APIs, and provide training that allows researchers to use them efficiently. Policy makers should ensure that data-intensive science requirements inform EuroHPC and EOSC priorities, enable multi-year allocation across heterogeneous resources, fund/reward sustained software development and data stewardship as first-class activities, support career pathways for research software and data engineers, and set environmental budgets that infrastructure operators are required to meet.

The Technical Blueprint (D6.1) provides the architectural foundation; the SRIDA provides the strategic framework for investment, governance, and implementation. Together they chart the path toward a 2035 compute and data continuum where researchers move workloads across HPC, HTC, cloud, and edge resources with one identity, predictable allocation, and portable software, and where environmental footprint is measured and managed. **The expected impact extends beyond the instruments that drove the agenda: a workforce skilled in research software and data engineering, European innovation in software, in advanced computing and energy-efficient hardware, broader public visibility of publicly funded science, and policy inputs on AI governance, digital sovereignty, and sustainable computing.**



# 1. Introduction



This section introduces the SPECTRUM Strategic Research, Innovation and Deployment Agenda (SRIDA), its relationship to the Technical Blueprint, and the audiences it serves. It establishes the scope, vision, and structure that frame the strategic priorities presented in subsequent sections.

## 1.1. Purpose and Scope

European data-intensive science faces a structural challenge. The instruments coming online over the next decade, the High Luminosity Large Hadron Collider, the SKA Observatory, LOFAR 2.0, and facilities across multiple domains, will generate data at scales that exceed current infrastructure capacity by an order of magnitude. Meeting this challenge requires more than incremental growth. It requires coordinated European action on governance, architecture, software, sustainability, and workforce.

This document presents the SPECTRUM SRIDA. It defines at an EU-level strategic priorities, investment areas, and a multi-annual roadmap for European research computing and data infrastructure. The agenda complements and builds upon existing European strategies: EuroHPC for supercomputing capability, EOSC for federated data and open science, and ESFRI for long-term research infrastructure planning. Where those strategies address broad European digital capacity, the SRIDA focuses specifically on the compute and data continuum required for data-intensive science, grounded in documented requirements from scientific communities. The network underpinning the edge and continuum is another important factor, however it is not treated in detail here and would require a dedicated focus in follow up workSPECTRUM (Computing Strategy for Data-intensive Science Infrastructures in Europe) brings together research infrastructures, digital infrastructure providers, and policy stakeholders under Horizon Europe. The project has gathered evidence through systematic use case analysis (D5.1), landscape survey (D5.3), access policy analysis (D5.2), community consultation via the Community of Practice (D3.1, D4.1), and expert review. This evidence base informs both the SRIDA and the companion Technical Blueprint (D6.1).

### Relationship with the Technical Blueprint

The SRIDA works in tandem with the Technical Blueprint. The Blueprint defines the architectural vision and technical capabilities for the compute and data continuum, identifying sixteen consolidated gaps. The SRIDA builds on these technical requirements, expands into non-technical areas  and formulate strategic actions: it explains why these capabilities matter for European research competitiveness, identifies who should lead implementation, proposes when investments should be made, and recommends where resources should be directed.

### Scope

The SRIDA focuses on infrastructures supporting data-intensive scientific research. Two flagship domains anchor the analysis:

- High Energy Physics (HEP), including the HL-LHC programme and the Worldwide LHC Computing Grid (WLCG), generating tens of exabytes of data from 2030 onwards and requiring distributed processing across heterogeneous compute resources at hundreds of sites globally
- Radio Astronomy (RA), including the SKA Observatory, LOFAR 1.0 and LOFAR 2.0, with data rates scaling to exabyte levels during full operations

### Scope and Transferability

SPECTRUM's priorities are grounded primarily in the documented needs of HEP and RA, complemented by additional science cases from life sciences, neuroscience, and meteorology collected in D5.1. Many of these needs are shared across data-intensive science, but the SRIDA does not presume to speak for communities that were not directly consulted. To enable other communities to assess which findings apply to their context, the following characteristics define the infrastructure assumptions underlying this agenda:

- Centrally reconstructed data: experimental data is correlated and/or reconstructed at central facilities (CERN Tier-0, SKA correlator) before it is distributed for further processing and analysis
- Research data, not personal data: primary datasets are physical measurements with no GDPR constraints; infrastructure operations (login, accounting) handle personal data separately
- Open science by default: data is intended for broad sharing with time-limited embargoes, there are no permanent access restrictions
- Throughput over latency: processing is optimised for high volume throughput (batch processing and real-time filtering at source) and not for low-latency delivery of results
- Global federated governance: experiments span multiple countries with formal governance structures
- Multi-decade experiment lifecycles: instruments operate for decades with forecastable data growth (HL-LHC schedule, SKA phases, LOFAR phases)

Many priorities in this agenda apply broadly across data-intensive science. Workforce capacity, energy efficiency, software sustainability, federated authentication, and long-term data preservation are challenges shared with life sciences, environmental monitoring, earth observation, and other domains. Where contexts differ, adaptation is needed: life sciences operates under GDPR for primary data; IoT and environmental monitoring work from independent streams without central correlation; real-time earth observation prioritises latency over throughput. In the following, the agenda distinguishes between domain-specific findings grounded mainly in HEP and RA evidence (see section 4, sub-section "Challenge" and Annex A) and cross-cutting recommendations applicable to data-intensive science more broadly (see section 4, sub-section "Recommendations").

## 1.2. Vision and Strategic Goals

### Long-Term Vision

"By 2035, European researchers will access a seamlessly integrated compute and data continuum spanning High Performance Computing, High Throughput Computing, cloud resources, quantum computing, and edge infrastructure. Scientists will execute complex workflows across heterogeneous resources without concern for infrastructure boundaries, authentication barriers, or data locality constraints. Europe will maintain digital sovereignty over its research data whilst participating in global scientific collaborations."

Realising this vision requires progress across several dimensions. Researchers need unified access to European computing and data resources through federated governance, harmonised allocation mechanisms, and standardised interfaces. FAIR data management must enable exabyte-scale workflows with automated placement, transport, and long-term preservation. Portable and reproducible workflows require distributable software stacks and architecture-optimized code that exploits heterogeneous hardware. Embedded AI/ML capabilities must support simulation, reconstruction, and analysis throughout scientific workflows. Operations must balance scientific ambition with environmental responsibility through green computing practices. Federated identity must ensure cybersecurity and European digital sovereignty across all resources.

### Strategic Goals

The SRIDA pursues the following five overarching goals, which structure the priorities presented in Section 4:

1. Coherent governance connects vertical, thematic research infrastructures with horizontal e-infrastructures through formal coordination mechanisms;
2. Seamless resource access enables researchers to work across infrastructure boundaries through multi-year allocation for RIs, aligned with evolution of their facilities, and federated authentication;
3. Technical interoperability standardises interfaces for heterogeneous computing and cross-facility workflow orchestration;
4. Sustainability ensures software-wise that codes remain portable across architectures and reproducible across decades-long programmes and environmentally-speaking that tools and policies are defined to keep impacts under control;
5. Human capacity develops expertise combining domain science with computing skills, supported by stable career pathways.



## 1.3. Target Audience

The SRIDA addresses three primary stakeholder groups whose coordinated action is essential for realising the compute and data continuum.

Scientific communities and research infrastructures include domain scientists, software developers, workflow engineers, and operations teams conducting data-intensive research. For this audience, the SRIDA provides a strategic framework and investment priorities (see Section 5) for digital infrastructure evolution aligned with published scientific roadmaps (e.g. the HL-LHC programme plan, the SKA Construction and Science Plan, and the LOFAR 2.0 plan), helping communities articulate requirements and engage with infrastructure planning.

Service providers and e-infrastructures include national computing and data centres, data centers from research institutions, EuroHPC sites, network operators (e.g. GÉANT), the EGI federation and the EOSC ecosystem. For this audience, the SRIDA offers strategic direction for service development and investment priorities, identifying where interoperability and coordination can enhance service delivery.

Policy makers and funding bodies include ESFRI delegates, the EuroHPC Governing Board, the European Commission, and national research councils. For this audience, the SRIDA provides evidence-based recommendations for strategic investment, grounded in documented requirements from scientific communities.

The SRIDA may also be of interest to secondary audiences including independent researchers outside large collaborations, who benefit from accessible entry points to shared infrastructure; universities, industry, and SMEs developing technologies or collaborating with research infrastructures; and, where relevant, citizens engaged with research infrastructures through public-engagement channels operated by the RIs themselves.

## 1.4. Document Structure

The SRIDA is organised as follows:

- Section 2 – Integrated Strategic Foresight Framework and Methodology: Strategic foresight framework and evidence gathering approach
- Section 3 – Strategic Context and Trends: European landscape and five transformations shaping infrastructure evolution
- Section 4 – Strategic Priorities: Thirteen priorities organised within the Four Pillars framework, with challenge analysis and recommendations for each pillar
- Section 5 – Investment Areas: Resource requirements mapped to priorities and investment types
- Section 6 – Multi-Annual Roadmap: Implementation phases across short-term, medium-term, and long-term horizons
- Section 7 – Broader Impact: Brief analysis of socio-economic impact
- Section 8 – Conclusion: Brief summary of the document
- Annexes: Detailed priority specifications, priority template and summary of supporting evidence

# 2. Integrated Strategic Foresight Framework and Methodology



## 2.1. Purpose and Rationale

The development of SPECTRUM's Strategic Research, Innovation and Deployment Agenda (SRIDA) and Technical Blueprint employs an integrated strategic foresight framework that builds upon the foundational work completed in WP3/WP4 (Community of Practice establishment) and WP5 (use cases, landscape analysis, and gap identification). This framework transforms the comprehensive analytical base into forward-looking strategies for data intensive European research infrastructure, informed by HEP and RA, over the next years. Rather than developing multiple scenarios or alternative architectures, SPECTRUM's approach focuses on creating an agreed vision that can adapt to different rates of technological change and varying implementation conditions. The framework employs foresight methods not to create competing futures, but to stress-test and refine a coherent strategy that addresses identified priorities whilst maintaining flexibility for evolution.


Figure 2.1.1: Integrated Strategic Foresight Framework Diagram.

## 2.2. Relationship to Technical Blueprint

The SRIDA and Technical Blueprint are complementary elements of a unified strategy. The Blueprint provides the technical architecture that enables achievement of the strategic priorities, whilst the SRIDA ensures that investments, governance, and implementation sequencing support the Blueprint's realisation.

Both documents were built upon substantial preparatory work. The Community of Practice (launched under WP3, and subsequently sustained and further developed through WP4) has established collaborations with over 50 participating entities, creating working groups across domains and technical areas that provide the distributed expertise essential for strategy development. It has also enabled the project to carry out a series of consultations and interviews with key stakeholders. The landscape analysis (WP5) has delivered comprehensive use case analysis, infrastructure mapping, and gap identification that grounds forward-looking strategies in current realities and validated requirements, all with the support of a Community of Practice (WP3)..

The relationship between documents serves different audiences. Technical implementers read the Blueprint for architecture then consult SRIDA priorities for stakeholder coordination and investment sequencing. Policymakers read the SRIDA for strategic rationale then reference the Blueprint for technical validation. Research communities use both documents together to understand what digital infrastructure can provide and how to influence its development.

## 2.3. Strategic Development Approach

The SRIDA development follows three integrated phases that move from analysis through synthesis to implementation planning.

The first phase, Anticipate, synthesises WP5 findings with forward-looking analysis to establish strategic context. This involves examining how identified gaps might evolve, projecting technology maturation timelines, and assessing how external factors (e.g., policy changes, funding landscapes, technological breakthroughs, environmental impacts) could influence implementation. Through Community of Practice (CoP) engagement, SPECTRUM validates and expands the vision for data intensive European research infrastructure, establishing the parameters within which both the Technical Blueprint and strategic priorities must operate.

The second phase, Analyse, develops and evaluates strategic components in parallel tracks. The Technical Blueprint track translates requirements into a coherent, high-level architecture design that addresses identified gaps whilst maintaining flexibility for technology evolution. Simultaneously, the SRIDA track analyses and prioritises strategic initiatives across governance, architecture, software, and workforce dimensions. Cross-impact analysis ensures technical feasibility aligns with strategic priorities, whilst the Community of Practice validates both architectural approach and priority rankings through structured consultation.

The third phase, Prepare, integrates the Technical Blueprint architecture with SRIDA priorities into a coherent implementation pathway (i.e., a roadmap). Using backcasting from the target state, SPECTRUM identifies critical milestones, resource requirements, and decision points. The Community of Practice transitions to co-designers and advocates, helping define implementation sequences and becoming champions for adoption within their organisations.

Throughout this process, evidence is drawn from internal sources (SPECTRUM project deliverables, Technical Blueprint requirements, CoP working group outputs) and external sources (European and international roadmaps from external sources, categorised in Annex C.1, including policy and funding bodies, strategic coordination groups, e-infrastructure providers, and thematic research facilities stakeholder interviews, and interactive sessions at conferences and workshops). Engagement with the three primary stakeholder groups (scientific communities and research infrastructures, computing and data service providers, and policy makers and funding bodies) occurs through technical workshops, CoP sessions, consultations, and the SPECTRUM Knowledge Hub.

Each priority undergoes systematic assessment across strategic importance, implementation complexity, stakeholder engagement, and measurability. The standardised priority template (Annex B) ensures consistent treatment across all thirteen priorities, with explicit traceability from identified gaps through strategic interventions to technical implementation. This methodology enables systematic comparison across priorities and provides clear accountability frameworks for implementation.



# 3. Strategic Context and Trends

European data-intensive science is entering a period of structural change. The instruments coming online over the next decade will generate data at scales that exceed current infrastructure capacity by an order of magnitude and more . Addressing this challenge requires coordinated European action across governance, technology, sustainability, and workforce dimensions. This section identifies the key drivers shaping this environment and their implications for the European research infrastructure strategy and the data intensive sciences with a focus on HEP and RA .

## 3.1. European Context

SPECTRUM operates within a European research and digital infrastructure ecosystem in which policy frameworks, coordination bodies, federations, sites and the scientific instruments they serve interact across multiple levels. Understanding this ecosystem is essential for positioning the SRIDA's priorities and identifying where its recommendations complement, extend, or depend upon existing initiatives. Figure 3.2 provides an overview, structured across five organisational levels (L0 scientific instruments, L1 sites, L2 federations, L3 coordination, L4 policy and regulation) and two tracks: a thematic track that is domain-specific and follows a single scientific community, and a horizontal (or digital) track that delivers cross-domain compute, data and network services. The remainder of this section describes each layer from top (L4) to bottom (L0), and closes with a note on the evolving infrastructure categories that intersect this picture.

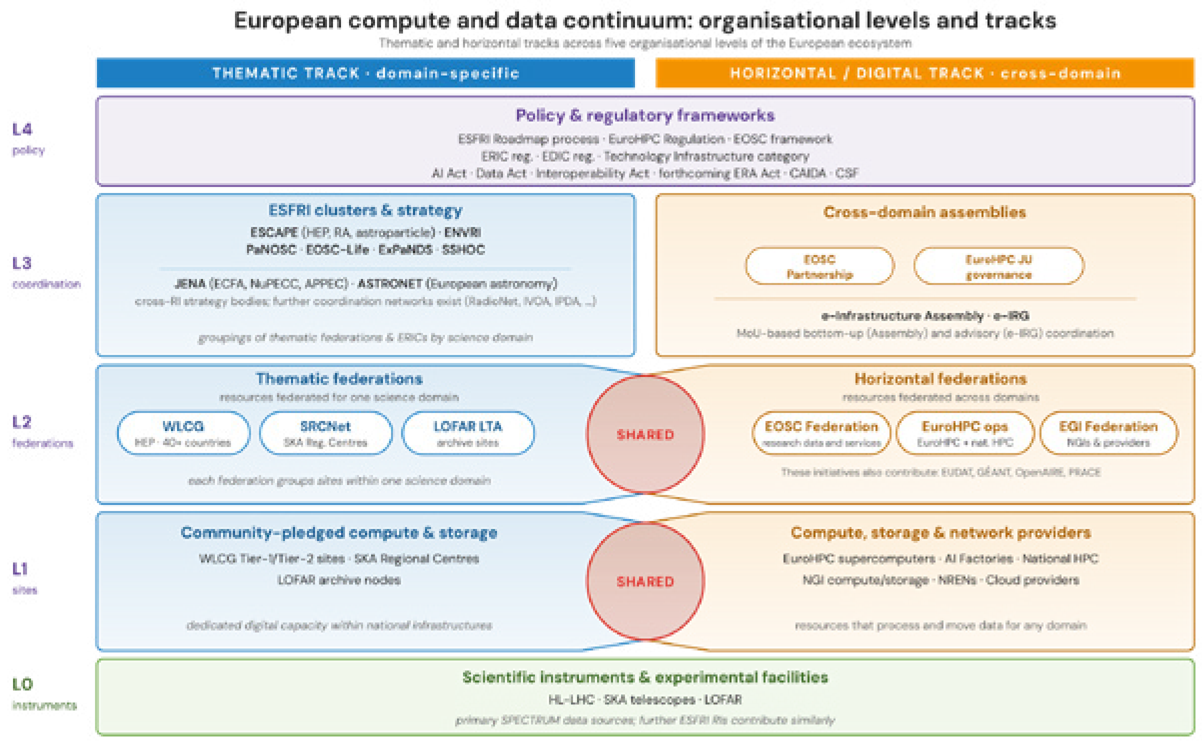



Figure 3.2: European compute and data continuum (organisational levels and tracks)

### (L4) Policy and regulatory frameworks

European policy frameworks provide strategic direction, governance and funding for research infrastructure at a European level. This matters because research, which is inherently cross-institutional and trans-national, requires coordinated decisions that single Member States cannot take alone. The EuroHPC Joint Undertaking, established under the EuroHPC Regulation, coordinates European supercomputing on a 50/50 co-funding model between the European Commission and participating states, with exascale systems becoming operational in 2025-2026. The European Open Science Cloud (EOSC) establishes the framework for federated data infrastructure, FAIR practices, and open science services; its Federation and EU Node are under active development. ESFRI, through its Roadmap and Landscape Analysis, has long recognised the strategic importance of the research infrastructures that anchor this agenda: the LHC programme (with HL-LHC operations extending to the late 2030s), the SKA Observatory (with full operations continuing beyond 2040), and LOFAR (operating since 2010, with LOFAR 2.0 extending operations into the next decade). RAISE (Resource for AI Science in Europe), launched in 2025, coordinates AI capabilities for science across the EU.

The regulatory and policy environment is evolving. The AI Act, Data Act, Interoperability Act, the forthcoming ERA (European Research Area) Act, and ongoing Omnibus developments establish new requirements for research infrastructure operations covering AI governance, data sharing, interoperability standards, and cross-border resource circulation. On the sovereign-supply dimension, the proposed Cloud and AI Development Act (CAIDA) aims to develop European cloud and AI infrastructure, complemented by the Cloud Sovereignty Framework (CSF), used by the Commission since April 2026 to grade providers across sovereignty criteria for public-sector procurement. The European Commission's 2025 Strategy on Research and Technology Infrastructures identifies research infrastructures (both the scientific facilities and the digital infrastructures that operate them) as underpinning European technological and data sovereignty. The 2028-2034 Multiannual Financial Framework allocates sustained investment through the European Competitiveness Fund, creating a funding pathway from research to deployment of digital research infrastructures.

### (L3) Coordination

Two complementary coordination structures organise European research infrastructure collaboration. The thematic track brings together research infrastructures within scientific domains. The horizontal track federates the cross-cutting compute, data and networking services on which all scientific domains depend.

Thematic coordination. The ESFRI Cluster projects, launched from 2019 under H2020 INFRAEOSC-05 and named as such in the ESFRI Roadmap 2021, connect research infrastructures listed on the ESFRI Roadmap around shared challenges in open science, data management and computing, with explicit links to the European Open Science Cloud. Five clusters cover astronomy, particle physics and astroparticle physics (ESCAPE), environmental science (ENVRI-FAIR), photon and neutron science (PaNOSC), life sciences (EOSC-Life) and social sciences and humanities (SSHOC); a sixth cluster for digital research infrastructures is in preparation. Cross-RI strategy is coordinated by JENA (Joint ECFA-NuPECC-APPEC) for particle, nuclear and astroparticle physics, and by ASTRONET for European astronomy. Further coordination networks operate alongside, including astronomy-specific networks (RadioNet, Europlanet, Opticon, Solarnet) and international standards alliances (IVOA, IHDEA and IPDA for virtual observatory, heliophysics and planetary data coordination respectively).

Horizontal coordination. On the digital infrastructure side, two tiers operate side by side. The formal EC-driven tier comprises the EOSC Partnership, the Horizon Europe co-programmed partnership that operationalises the EOSC framework, and the EuroHPC JU governance; both coordinate top-down with EC policy and funding. A bottom-up advisory tier comprises the e-Infrastructure Assembly, an MoU-based forum where the major European e-infrastructures (EGI, GÉANT, PRACE, EUDAT and others) coordinate operationally, and the e-Infrastructure Reflection Group (e-IRG), a Member-State-level advisory forum that aligns digital research infrastructure policy with national programmes. OpenAIRE coordinates open-access scholarly infrastructure, EVERSE addresses research software quality and sustainability, and European Data Spaces establish frameworks for cross-sectoral data sharing.

### (L2) Federations

The federation layer pools resources from many sites into a single operational service that researchers can use without needing to interact with each site individually.

Thematic federations pool sites within one scientific community. The Worldwide LHC Computing Grid (WLCG) coordinates federated computing and data resources across more than 40 countries for the LHC experiments, operating under a multi-party memorandum of understanding (MoU) governance model. The SKA Regional Centres Network (SRCNet) is being designed from inception as a federated network of national and regional centres serving SKA data to end users and supporting their analysis needs. The LOFAR Long-Term Archive (LTA) coordinates distributed data processing across European LOFAR stations and a federated data archive.

Horizontal federations pool sites across many scientific domains. The EGI Federation aggregates compute and storage delivered by National Grid Initiatives (NGIs) and other providers. GÉANT federates the pan-European networking backbone delivered by the National Research and Education Networks (NRENs). The EuroHPC operations layer, consisting of EuroHPC-procured supercomputers and the national HPC centres that operate them, functions



as the horizontal compute federation parallel to EGI and GÉANT. PRACE (Partnership for Advanced Computing in Europe) operates a continuing HPC partnership association alongside EuroHPC, with its own access calls and community programmes. The EUDAT Collaborative Data Infrastructure (CDI) federates more than 20 European research organisations and data centres delivering the B2-suite data services, and is on a path to join the EOSC Federation as a Node, parallel to EGI's onboarding.

### (L1) Sites

The site layer is where compute, storage and network capacity is physically delivered.

Thematic sites carry community-pledged capacity dedicated to specific science domains. For HEP these are the WLCG Tier-0 (CERN), Tier-1 (currently 14 national centres) and Tier-2 (around 150 institutional centres) sites. For SKA these are the SKA Regional Centres in their national or regional implementations. For LOFAR these are archive nodes that host the federated long-term archive.

Horizontal sites carry capacity that processes and moves data for any domain. EuroHPC-procured supercomputers (such as LUMI, LEONARDO and MareNostrum 5), hosted by national consortia under EuroHPC governance, deliver European-level high-performance capability. National HPC centres, funded predominantly through national programmes, provide the bulk of additional HPC capacity. NGIs (national e-Infrastructures) and other compute providers deliver the resources federated by EGI. NRENs deliver the network capacity that GÉANT connects at European level. Commercial cloud providers supply elastic capacity where research workflows allow. The EuroHPC AI Factories, becoming operational from 2025-2026, provide AI-optimised compute alongside national HPC.

### (L0) Scientific instruments

The instrument layer is what the rest of the ecosystem ultimately exists to serve. For SPECTRUM, the primary instruments are the High-Luminosity LHC at CERN, the SKA telescopes operated by SKAO, and the LOFAR telescope operated by ASTRON. Further ESFRI research infrastructures generate comparable scientific data and contribute to the broader continuum but are not the SRIDA's primary focus.

### Evolving infrastructure categories

The European infrastructure landscape is expanding beyond the traditional Research Infrastructure model. The European Commission is introducing new infrastructure categories and pushing towards a more integrated landscape in which research, technology, and digital infrastructures complement one another. Three legal-instrument categories now coexist:

- Research Infrastructures (ESFRI Roadmap, ERIC regulation): the established pathway for publicly funded scientific facilities. CERN, SKAO and LOFAR ERIC operate under this model. The ESFRI Roadmap process signals strategic importance and facilitates engagement with national ministries.
- Technology Infrastructures (TIs): a category promoted by the EC, focused on facilities for developing, testing, and validating technology to foster industrial competitiveness. TIs are oriented towards the commercial sector, with the expectation that industry pays for services. The RI/TI distinction is one of primary mission and beneficiary (research community vs industrial sector) rather than of operational maturity: many RIs, including the LHC programme and SKA Observatory, operate at very high TRL.
- Digital Infrastructures: distinct from RIs in legal form (the EDIC regulation, under the Digital Decade, provides the cross-border legal instrument, parallel to the ERIC regulation for RIs) and in primary orientation (Digital Decade goals for public and commercial digital transformation).

The EC is actively promoting integration across these categories, seeking to avoid silos and build a continuum of "knowledge infrastructures" spanning research, technology, digital, and data dimensions. For SPECTRUM, this matters in two ways. First, the research-IT compute and data continuum serves publicly funded research communities. Historically, digital research infrastructures have been funded through Horizon work programmes rather than the ESFRI Roadmap, but ESFRI's recognition of EOSC and the rise of the DIGIT working group are moving them into closer alignment with the RI tradition. Second, the technologies and services this continuum develops (federated computing, data management, AI platforms) have broader applicability that the EC's integrated infrastructure vision recognises. Engagement with these emerging categories, particularly where the EC encourages collaboration between RIs and TIs, is an active question for positioning the agenda's priorities.

### Foundation for this agenda

SPECTRUM builds on two decades of operational experience from the federations and sites described above. National funding provides the majority of computing capacity across all of them: EuroHPC operates on a 50/50 co-funding model, while the resources underpinning WLCG, SRCNet, LOFAR and EGI are predominantly nationally funded. Within EOSC, an in-kind co-funding model is being trialled as a potential mechanism for sustaining pan-European services, complementing direct national and EU funding streams. The documented requirements and identified gaps from these communities, combined with analysis of European strategy documents, inform the strategic priorities presented in Section 4. For detailed descriptions of European initiatives, see D6.1 Technical Blueprint Annex B.

## 3.2. The Evolving Landscape

Seven external drivers are reshaping the strategic environment for European data-intensive science. These drivers emerge from the landscape analysis within WP5, of scientific use cases (Deliverable 5.1), infrastructure capabilities (Deliverable 5.3), access policies (Deliverable 5.2), community consultation, and alignment with European strategy documents. Each driver represents a force that current infrastructure arrangements cannot adequately address through incremental change alone. Together, they define the transformation that the strategic priorities in Section 4 are designed to deliver.


Figure 3.2.1: Seven strategic drivers shaping the European compute and data continuum.

### 3.2.1. Digital Infrastructure Scale

European flagship instruments will generate data volumes that exceed current processing, storage, and networking capacity by an order of magnitude from 2030 onwards. Both flagship programmes are projected to generate hundreds of petabytes per year at full operation (HL-LHC during Run 4-6, SKA1 from AA4 onwards), with multi-exabyte aggregates over their lifetimes; SKA's challenge is concentrated in the real-time signal path, HL-LHC's in long-term storage and reprocessing. Meeting this scale through linear growth alone (e.g., adding nodes, disks, and bandwidth) would breach power, cost and procurement envelopes by the early 2030s. Architectural change is therefore required across compute, storage, and networking (heterogeneous computing, near-data processing, federated storage), as set out in detail in the Technical Blueprint (D6.1 Section 3).

Meeting this scale requires architectural change across compute, storage, and networking (heterogeneous computing, near-data processing, federated storage), not linear scaling. Two factors compound the pressure. On the supply side, efficiency gains from new hardware tend to be absorbed by increased demand (the rebound effect documented in D6.1 §3.2), and in the current European energy mix the fabrication-phase environmental footprint of replacing hardware can outweigh usage-phase savings. On the demand side, AI workloads (§3.2.2) increasingly share the same infrastructure as data-intensive science, pulling resources in directions that traditional simulation does not. The architectural alternatives, scheduling and life-cycle practices needed are set out in the Technical Blueprint (D6.1 §3.2, §4 capabilities and §5.3 software portability).

European exascale systems offer new and transformational computational capabilities, but realising their potential for science requires addressing a fundamental mismatch: strategically important data intensive research infrastructures (e.g., ESFRI roadmap) operate on multi-decade timelines, yet access to HPC resources typically follows annual competitive cycles. Major experiments need planning certainty over 5-10 year horizons; current competitive, rather than strategic, allocation models provide 1-2 years at most. Bridging this temporal gap is essential for effective (digital and research) infrastructure utilisation, planning and alignment.

Architecture diversity adds complexity. Computing platforms now span multiple processor families and accelerator types, each requiring distinct software optimisation. Scientific workflows must operate across this heterogeneous landscape. Investment in portable software infrastructure (e.g., EESSI - the European Environment for Scientific Software Installations - and containers) can reduce fragmentation, but requires coordinated community effort, in-depth hardware architecture skills and sustained funding. In addition, the European systems need to consider high-volume data processing throughput as an essential criterium in the design of these systems. This means optimizing the full chain from data repository access (connecting to distributed storage systems), networking (connecting datacentres) and the system configuration within the datacentre (e.g., network, staging/processing storage, compute server layout).

Strategic implication: Europe needs allocation mechanisms that bridge the mismatch between infrastructure timelines and funding cycles, alongside investment in software portability that enables efficient use of diverse architectures. High-volume data processing throughput needs to be an essential criterion in the



design of European systems, and these systems need to be connected through federated identity, interoperable interfaces and unified data movement so that researchers can compose pipelines across HPC, HTC and cloud resources without per-stack glue. Strategic planning, allocations and co-design with the data-intensive research communities is essential for achieving this.

### 3.2.2. AI Adoption

Artificial intelligence is being adopted rapidly across the research data lifecycle as a tool that supports simulation, reconstruction, analysis, and operations. Adoption is uneven and the scientific value of AI depends on the circumspection with which it is used, but demand for AI-capable infrastructures is growing fast across European research communities. Examples already in production span across scientific  domains: AlphaFold for protein structure prediction, ECMWF machine-learning weather models that rival physics-based simulation for medium-range forecasting, and CNN-based gravitational-wave signal detection in LIGO/Virgo. The pattern is consistent: AI accelerates specific stages of the research workflow rather than replacing the scientific method that surrounds them.

A maturity gap constrains progress. Most AI applications in physics remain at the proof-of-concept stage rather than production deployment. Scaling from demonstration to operational use is challenging, often dependent on small teams rather than sustainable capacity. The gap between AI potential and AI reality represents a strategic bottleneck.

Hardware evolution creates both opportunity and tension. The computing industry is optimising for AI workloads that differ from traditional scientific simulation, creating divergence between commercial hardware roadmaps and scientific computing needs. Some workloads (e.g., AI inference, signal processing, parts of generative simulation) can exploit reduced-precision AI-optimised hardware directly; others (e.g., Lattice QCD, climate, high-precision lattice models) remain FP64-bound and risk diminishing alignment with mainstream technology development unless mixed-precision strategies or dedicated capability systems are sustained alongside AI-optimised platforms.

Scientific communities must either adapt to AI-optimised hardware or accept diminishing alignment with mainstream technology development.

Alternative approaches are emerging. Smaller, more efficient AI models can run on distributed infrastructure closer to data sources, reducing the need for massive centralised resources. Domain-specific scientific AI models, developed by and for research communities, provide capabilities tailored to scientific needs rather than commercial applications, but their training, validation, and continued availability depend on curated scientific datasets, dedicated training compute, and governance arrangements that only sustained research infrastructure can underwrite.

European policy is also shaping AI ambition. The AI Continent Action Plan commits substantial public investment to AI infrastructure (over €10 billion for AI Factories, €20 billion for AI Gigafactories, and a target of tripling European data-centre capacity), and the AI Act sets the regulatory baseline for trustworthy deployment. For research specifically, RAISE (Resource for AI Science in Europe), launched in November 2025, coordinates AI resources along two dimensions, "Science for AI" and "AI in Science", through a Secretariat, Digital Hub, and Academy; its first project, SCIANCE, is mapping the European AI-in-science landscape. The direction of AI in research is no longer set by demand and technology alone, it is increasingly shaped by these European policy choices.

Strategic implication: Europe needs community-driven AI infrastructures that serve scientific requirements. Governance arrangements, including coordination instruments such as RAISE, must let research communities shape AI deployment rather than reduce them to consumers of commercial offerings.

### 3.2.3. Environmental Sustainability

Energy availability and environmental impact are becoming primary constraints on infrastructure growth. Technically speaking, some locations face grid and cooling capacity limits. Politically speaking, in a context of multiple environmental crises, societal expectations for environmental responsibility are rising. Meanwhile, the use and the energy intensity of AI and large-scale computing are increasing rapidly, so that data centre power demand and environmental impact are projected to grow substantially over the coming decade,  overwhelmingly driven by commercial AI workloads. Research-IT growth, while non-trivial, remains a small share of that total, and European public research infrastructures lead on efficiency (see Green500 standing of EuroHPC systems).

Environmental accountability is becoming a condition of public funding and social license to operate. To resolve this tension, hardware and software tools have emerged to measure and report resource usage and methodologies based on life-cycle assessments and even ISO standards defining systems and metrics are being developed to quantify environmental impact. However, this is done piecemeal and a posteriori, instead of globally and proactively in order to contain impacts and avoid transfers Europe has demonstrated leadership in energy-efficient high-performance computing. European systems achieve world-leading efficiency metrics, as shown in EuroHPC supercomputers global standing in Green500 [85]. Emerging computing paradigms (photonics, neuromorphic architectures) offer potential for further efficiency gains. But realising this potential requires sustained investment and coordinated adoption. Sustainability also stands at the core of European policy initiatives — the Digital Decade and the European Green Deal frame the "twin transitions" (digital and green) introduced under the 2019-2024 Commission, and is a key driver of future European competitiveness.  Strategic implication: Europe needs a strategic, coordinated approach to sustainability in research computing, bringing digital infrastructures, research infrastructures, and the Commission together, supported by an EU-level action that catalyses the collaboration, common sustainability standards, accountability mechanisms covering the full facility life cycle, and investment in energy-efficient technologies and operational practices.

### 3.2.4. Security and Trust

The security landscape is evolving in ways that require proactive adaptation. Quantum computing developments create urgency for cryptographic transition. Human-based authentication requirements increasingly, due to single person failure, conflict with the automation that scientific workflows require as well as for collaborative asset-maintenance and the incorporation of AI like AI agents. Federated operations across organisational and national boundaries depend on several trust frameworks that in terms of cross-framework alignment require additional and coordinated effort at an EU-level.

Data governance is evolving. European Data Spaces initiatives are establishing standards for secure data sharing that research infrastructures must align with to ensure interoperability with the broader European regulatory frameworks (GDPR, Data Act, Data Governance Act, and AI Act). The scientific data and software produced by European research communities has important societal and commercial impacts and hence are unique assets requiring dedicated stewardship (at an EU-level) that cannot be outsourced. Therefore the constraints arisen from digital sovereignty drivers also have an impact on security and trust  Strategic implication: Europe needs proactive security evolution, federated trust frameworks enabling cross-boundary collaboration grounded in EU law and regulations whilst remaining fit for purpose, and alignment with emerging European data governance frameworks.

### 3.2.5. Long-Term Preservation

Multi-decade instrument timelines create preservation challenges that extend beyond immediate operational concerns. Scientific software and data must remain usable across hardware generations and organisational changes spanning at least 10-50 years. The expanding scope of the FAIR principles (now encompassing software, workflows, and AI models alongside data) reflects a growing recognition that reproducibility depends on preserving the complete digital context, not just datasets.

Active preservation requires sustained investment, and an explicit ownership model for the long-term cost of keeping data assets accessible, covering software functionality as underlying platforms evolve, data-format migration as standards change, and documentation sufficient for future reproduction. Software reuse should be encouraged via libraries usable across projects. Current funding models rarely support such long-term stewardship.

Strategic implication: Europe needs preservation strategies ensuring long-term (decades) reproducibility, sustained funding for software and data stewardship, and FAIR practices extended across the complete computational context. Re-use of data and reproducibility of processes are key elements for not only the scientific method but also for societal and commercial impact..

### 3.2.6. Workforce Capacity

The research computing workforce is undergoing fundamental change. In scientific domains, there used to be little distinction between researchers working on hardware, software and analysis tasks; this is shifting due to the larger specialisation in experimental physics, and with the advent of new technologies which need a more solid background in computing: skills requirements are shifting as heterogeneous architectures, machine learning methods, and security demands reshape what expertise is needed. Career structures designed for traditional academic paths do not accommodate the IT roles that modern research infrastructure requires. Competition with industry for technical talent has intensified as the same skills command premium salaries in the commercial sector. At the same time the research sector provides a unique and innovative frontier for digital challenges. Hence, the capability of upskilling workforce spilling over from academia to industry is also a driver for economic and social social impact Workforce capacity affects capacity across all other areas. Scale requires specialists in heterogeneous computing, software portability, and performance engineering. AI depends on specialised engineers who can bridge the gap between proof-of-concept and production deployment. Environmental sustainability goals require expertise in code optimisation, energy-aware scheduling, and efficiency monitoring. Trust and continuity demand security specialists and preservation experts who understand both technical and policy dimensions.

Europe faces a significant skills gap in precisely these areas. Community surveys indicate that the majority of research groups identify technical capacity as a concern, yet lack the resources to address it. Scientific career structures often do not recognise or reward contributions to infrastructure, software development, or operational efficiency. Research software engineers, who bridge domain science and computing expertise, typically lack defined career paths. This is further explained, for example, in the Briefing Book from the European Strategy Particle Physics update [143]. Incentivizing the creation and sustainability of IT skills together with the research community can be another important opportunity in the co-design process. Strategic implication: Europe needs coordinated career frameworks that recognise and reward research and digital infrastructure contributions, training programmes aligned with emerging skill requirements, and retention strategies that can compete with industry for technical talent.

### 3.2.7. Digital Sovereignty



European research infrastructure increasingly operates within a policy environment that treats digital capacity as a strategic asset. The European Commission's 2025 Strategy on Research and Technology Infrastructures identifies research infrastructures, both the scientific facilities and the digital infrastructures that operate them, as "underpinning European technological and data sovereignty". Frameworks alone are insufficient: sovereignty also depends on sustained European data infrastructure (storage capacity and the institutional expertise to curate it). The Cloud Sovereignty Framework (CSF v1.2.1, October 2025, EC DG DIGIT) defines eight Sovereignty Objectives (notably SOV-4 Operational Sovereignty, on the EU's practical ability to run and evolve technology independently) and a five-level SEAL assurance scale used in public-sector procurement. The 2028-2034 MFF allocates sustained investment through the European Competitiveness Fund, creating a funding pathway from research to European strategic autonomy. The strategic backdrop is quantified: a Commission study estimates a European compute capacity gap of around 20 GW by 2036, and three hyperscalers currently hold approximately 70% of the European cloud market.

For data-intensive science, sovereignty concerns manifest in three dimensions. First, dependence on non-European technology providers for critical components (processors, interconnects, cloud platforms) creates supply chain vulnerabilities that, for example, the EuroHPC Joint Undertaking addresses through open technologies such as RISC-V and the European Open Stack. Second, scientific data and software produced by European research communities are unique assets requiring processing within European jurisdictions (GDPR, Data Act, Data Governance Act, AI Act) under European governance, particularly as AI workflows (see §3.2.2 and P07) increasingly involve sensitive or strategically valuable datasets. Third, reliance on commercial AI models for scientific applications raises questions of reproducibility, auditability, and long-term availability that open-source, community-developed alternatives can address.

These concerns are not hypothetical. Recent international developments, including funding cuts, database closures, and restrictions on research collaboration, demonstrate that open science requires resilience frameworks capable of withstanding geopolitical disruption. European federated infrastructures, built on open-source technology stacks with multi-jurisdictional governance, provide this resilience by design. Europe's procurement response operates within tight legal constraints. Equal-treatment obligations (2019 EC guidance on third-country bidders) prevent excluding bidders on origin alone; available levers are the International Procurement Instrument (Regulation (EU) 2022/1031) for reciprocal restrictions, IPCEI for European industrial-capacity state aid, and CSF tender-grading within those rules. Stronger preferences likely require new legislation: the forthcoming Cloud and AI Development Act (CAIDA, target 27 May 2026, EU technological sovereignty package) targets efficient data centres, a tripling of EU data processing capacity by 2030, and sovereign EU solutions for 100% of highly critical use cases. The Berlin Declaration (November 2025, 17 Member States at ministerial level) anchors the political framing, including protection against extraterritorial application of non-EU laws. Strategic implication: Europe needs sovereignty-aware infrastructure strategies that ensure European control over critical research computing assets through open technologies, federated governance, and sustained investment in European digital capacity, while maintaining the openness and international collaboration essential to scientific progress and at the core of European digital norms and values (e.g., the European Declaration on Digital Rights and Principles).

## 3.3. From Drivers to Priorities

Responding to these drivers requires coordinated action across four dimensions: governance, technical architecture, workforce, and the business/funding models that sustain operation over the lifetime of the supporting infrastructures. No single driver maps to a single response: Infrastructure Scale demands both architectural evolution (heterogeneous computing, data management) and institutional change (multi-year allocation mechanisms). Digital Sovereignty requires open technology choices, federated governance models, and workforce investment in European capabilities. Environmental Sustainability spans infrastructure design, operational practices, and policy frameworks. This cross-cutting nature shapes the structure of this agenda. This agenda organises thirteen strategic priorities within four pillars, each addressing a distinct dimension of intervention. Pillar 1 (Policy, Trust and Governance) establishes the institutional frameworks enabling coordinated action, including cross-infrastructure governance, resource allocation mechanisms, and federated identity. Pillar 2 (Architecture and Interoperability) addresses the technical infrastructure for heterogeneous computing, data management, and workflow orchestration. Pillar 3 (Software and Science Enablers) develops the tools and practices that translate infrastructure into scientific capability, spanning AI integration, portability, reproducibility, and preservation. Pillar 4 (Human Capital and Responsibility) ensures the workforce capacity, community collaboration, and environmental sustainability foundations for long-term success. The pillars follow a sequencing logic reflected in the roadmap (Section 6). Governance and trust foundations (Pillar 1) must be established before technical integration can proceed at scale (Pillars 2 and 3), while workforce development and sustainability (Pillar 4) span all phases as enabling conditions. Priorities within each pillar are interdependent and reinforce each other; dependencies across pillars are addressed in Section 4 and the implementation roadmap. Section 4 presents these priorities with their challenges, evidence base, and recommendations. Annex A provides detailed specifications for each priority, including implementation pathways and success indicators.

# 4. Strategic Priorities



Building on the drivers and pillar framework introduced in Section 3.3, this section sets out the thirteen strategic priorities and their evidence base. Each priority addresses gaps identified through systematic analysis of scientific use cases, infrastructure capabilities, and community requirements, aligned with European strategy documents from EOSC, EuroHPC, ESFRI, and domain-specific roadmaps. Recommendations in section 4.x.3 follow the cross-cutting transferability convention introduced in section 1.1; HEP- and RA-specific evidence sits in the section 4.x.1 Challenge sub-sections.

## 4.1. The Four Pillars Framework

The thirteen priorities are organised into a Four Pillars framework. Each pillar represents a distinct dimension of intervention required to achieve the continuum vision for data intensive research, and together they form a coherent strategy where progress in one dimension enables and depends upon progress in others.

Pillar 1: Policy, Trust and Governance establishes the institutional foundations for coordinated infrastructure. This pillar addresses funding coordination mechanisms, multi-year resource allocation frameworks, and federated identity systems enabling authenticated access across organisational and national boundaries.

Pillar 2: Architecture and Interoperability creates the technical machinery for a coherent continuum. This pillar enables heterogeneous computing integration across heterogeneous computing (e.g., HPC, HTC, cloud, quantum and accelerated platforms); data federation and transport at exabyte scale; and workflow orchestration across distributed resources.

Pillar 3: Software and Science Enablers bridges infrastructure capabilities with scientific applications. This pillar advances AI (including ML/DL methods and emerging LLM/agentic approaches) and accelerated computing, software portability across architectures, reproducibility over multi-year timescales, and code preservation across the decades-long lifetimes of major research facilities (e.g., HL-LHC, SKA, LOFAR 2.0).

Pillar 4: Human Capital and Responsibility addresses organisational and environmental dimensions. This pillar builds collaborative communities through co-design processes, establishes environmental sustainability through impact-aware facility management, and develops career pathways combining domain science with computing expertise.

The pillar structure reflects a fundamental principle: all four dimensions must advance together. No single pillar succeeds in isolation. Research infrastructures cannot effectively access diverse computing resources (Pillar 1) without data management capabilities for exabyte-scale workflows (Pillar 2). Technical capabilities prove unsustainable without workforce development (Pillar 4) and lose scientific value without portability and reproducibility mechanisms (Pillar 3). All four pillars rest on a federated digital infrastructure that supplies compute, storage and network capacity through thematic and national centres, EuroHPC and federation initiatives such as EGI; sustaining this physical foundation is a precondition for any of the pillar interventions to deliver value. The diagram below illustrates the thirteen priorities organised within their respective pillars.

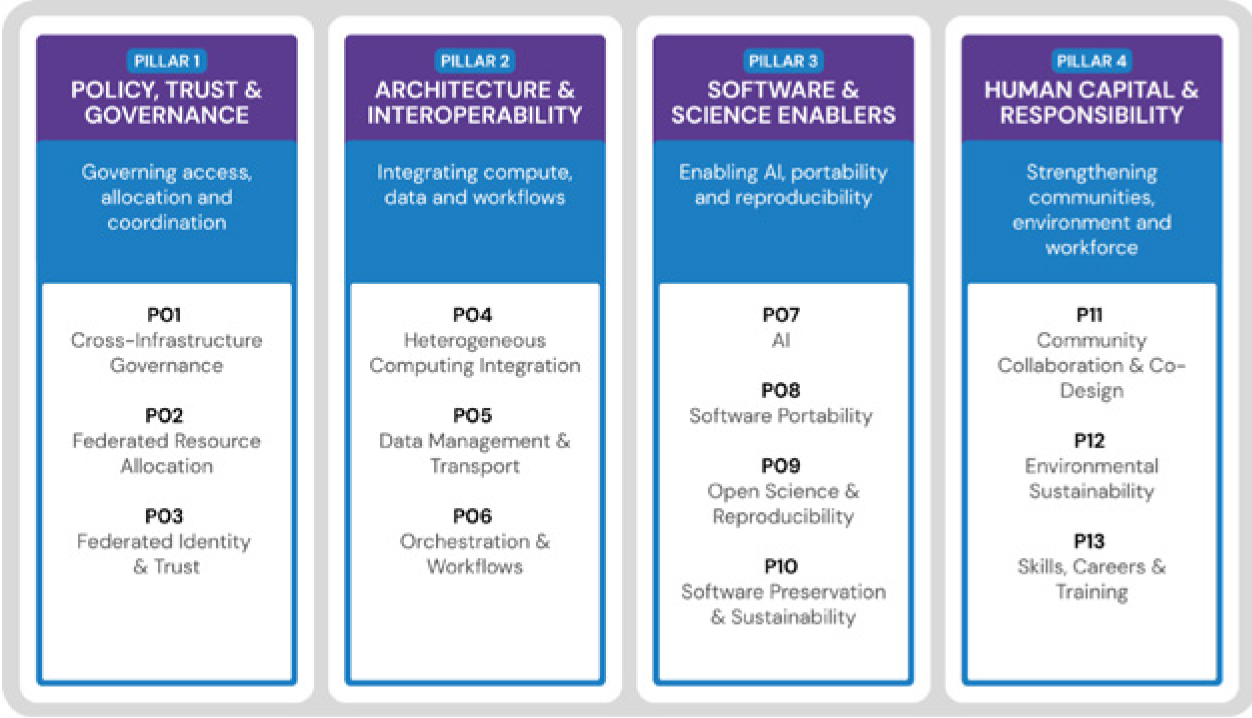



Figure 4.1.1: Strategic Pillars of the European Compute & Data Continuum for Research.

**Mapping to the Strategic Goals**
Each pillar realises one or more of the five overarching goals introduced in §1.2:

Table 4.1: Mapping SRIDA Goals to SRIDA Pillars

| Pillar | Goals from Section 1.2 |
|---|---|
| Pillar 1: Policy, Trust & Governance | Coherent governance; Seamless resource access |
| Pillar 2: Architecture & Interoperability | Technical interoperability |
| Pillar 3: Software & Science Enablers | Sustainability (software portability, reproducibility, preservation) |
| Pillar 4: Human Capital & Responsibility | Sustainability (environmental); Human capacity |

Sustainability spans Pillars 3 and 4 because its two sides, keeping codes portable and reproducible across decades-long programmes and keeping infrastructure environmentally responsible, are operationally distinct interventions delivered by different communities. Governance and access are grouped within Pillar 1 because both are institutional foundations, addressing the question of who coordinates resources and how researchers gain access to them, before the technical machinery of the continuum is introduced in Pillar 2.

# 4.2. Pillar 1: Policy, Trust & Governance

Governing access, allocation and coordination

## 4.2.1. Challenge

Europe invests in research computing and data infrastructure through several routes: EuroHPC JU is the principal EC-level channel for compute capacity; EOSC coordinates the federated data-services layer; the bulk of operational HTC, cloud and storage capacity is funded through national programmes and the thematic research infrastructures themselves; Horizon Europe (and successor FP10) work-programme lines support development and integration projects. Yet these investments face a structural challenge: computing and storage requirements are projected to grow by an order of magnitude by the 2030s, while EC-level capacity growth (channelled chiefly through EuroHPC and the AI Factories) is concentrated on exascale and AI-optimised systems, and national operational baselines that fund the bulk of HTC, storage and thematic-RI capacity are not projected to track the demand curve.

Moreover, researchers cannot easily use existing resources across organisational boundaries: the limiting factor is typically national or organisational funding eligibility rather than technical interoperability. EuroHPC and EOSC partially resolve this at EC level, but their access modes are not yet adapted to the long-running, data-heavy workflows characteristic of HEP and RA. A workflow requiring EuroHPC cycles, EGI federation services (which deliver capacity from national and thematic providers), and EOSC data repositories must navigate three separate governance structures, allocation processes, and authentication systems. This fragmentation reduces infrastructure utilisation and creates barriers that favour well-resourced teams over smaller research groups.

Three interconnected challenges prevent the compute and data continuum from functioning as coherent European capability:

- Governance fragmentation: Thematic research digital infrastructures (e.g. WLCG for HL-LHC, SRCNet for SKAO) and pan-European multi-disciplinary digital infrastructures (e.g. EuroHPC, EOSC, EGI) operate under different governance models with limited systematic coordination. Success stories exist: the long-standing cooperation between WLCG and EGI demonstrates that effective co-design is achievable. However, structured engagement between data-intensive science communities and EuroHPC remains underdeveloped. Wider coordination mechanisms would benefit all parties.
- Allocation cycle mismatch: Research infrastructures operate on multi-decade timelines (the LHC has been operational since 2008 and via the HL-LHC upgrade will run through 2041; SKA is foreseen to be operational for 50+ years from the start of operations), yet must navigate incompatible allocation models across the infrastructure landscape. Federated infrastructures like WLCG operate on pledged resources committed by national funding agencies, whilst EuroHPC centres use competitive peer-reviewed proposals typically granted for one year. Experiments need guaranteed access for 3-5 year periods compatible with their planning cycles, but no unified framework spans these different approaches. This creates planning uncertainty, repetitive submission burden, and discourages investment in workflow optimisation.
- Authentication fragmentation: federated digital infrastructures and HPC centres operate partially incompatible identity management systems. Lack of federated identity management and multi-factor authentication (MFA) barriers for unattended workflows prevent seamless authenticated access. Without this, technical interoperability alone provides no benefit to researchers.

## 4.2.2. Priorities Overview

Pillar 1 establishes governance, allocation, and trust foundations across both compute and data dimensions of the continuum. Three priorities form a layered intervention: governance coordination (P01) creates the institutional basis for allocation agreements (P02), which in turn require authenticated identity (P03) to track usage and enforce entitlements.

P01: Cross-Infrastructure Governance. Establish a coordination forum enabling thematic research digital infrastructures (e.g. WLCG, SRCNet, LOFAR ERIC) to align positions and seek formal representation within pan-European multi-disciplinary digital infrastructures (EuroHPC, EOSC, EGI), with policy-maker representation from national funding agencies and relevant EC units. The forum addresses interoperability across legal, organisational, semantic, and technical layers, coordinates an innovation agenda for data-intensive science distinct from classical HPC optimisation, and develops policy recommendations bridging operational versus project-based funding. National programmes provide the majority of HTC, cloud, and thematic-RI storage capacity, while the EC channel via EuroHPC concentrates on exascale and AI-optimised compute; this division of labour is deliberate, not legacy, and the strategic question is alignment across the two funding streams (consistent with the forthcoming ERA Act framework), not migration of one into the other. (Detail: Annex A, P01)

P02: Federated Resource Allocation. Establish multi-year allocation mechanisms aligned with timelines on which research facilities evolve, supported by standardised, transparent accounting frameworks for heterogeneous resources (CPU, GPU, storage, network) across governance boundaries. Existing models demonstrate feasibility: WLCG enables 5-10 year resource planning through national commitments, and EuroHPC's Strategic Access designation provides a pathway for enabling guaranteed multi-year allocations at the EU-level. Funding bodies can quantify return on investment through standardised accounting metrics tied to multi-year allocation checkpoints, addressing the operational-versus-project funding mismatch that currently undermines long-term scientific programmes and their societal impact. (Detail: Annex A, P02)

P03: Federated Identity and Trust. Deploy federated authentication enabling researchers to access infrastructure across organisational boundaries, with dynamic authorisation, automated delegation, and service-account support for long-running and unattended scientific workflows. The major community and federation AAI proxies, including Indigo-IAM (WLCG, SRCNet), EGI Check-in (EGI), MyAccessID (EuroHPC, EOSC, GEANT) are already AARC-BPA-compliant and accept eduGAIN identities; the integration challenge is interoperation across these proxies (attribute mapping, group and role assertion handling, IdP discovery), together with bridging to site-local SSO systems (e.g. CERN-SSO) for resources under separate authorisation domains. Adopt the AARC BPA as the architectural baseline, contribute to AARC-TREE work on cross-proxy attribute exchange, and make collaborative and service-account operation a first-class capability with delegation patterns that displace the single-person MFA workaround for unattended workflows (Detail: Annex A, P03).

Interdependencies: the three priorities are mutually load-bearing. Allocation (P02) cannot operate without governance agreements (P01) to legitimise cross-boundary use, and depends on federated identity (P03) to attribute usage; trust frameworks (P03) require P01 to gain currency outside their home organisation; and P01's coordination work becomes operationally meaningful only when P02 and P03 are in place.

## 4.2.3. Recommendations

For Scientific Communities and Research Infrastructures:

- Seek formal representation in horizontal e-infrastructure advisory structures to ensure scientific requirements shape infrastructure decisions (P01)
- Publish multi-year resource plans aligned with experimental timelines to inform infrastructure planning (P02)
- Adopt federated identity frameworks enabling seamless access across infrastructure boundaries (P03)

For e-Infrastructures and Service Providers:

- Establish liaison mechanisms with thematic research infrastructure governance bodies to ensure scientific requirements shape infrastructure decisions (P01)
- Technically implement multi-year allocation frameworks providing planning certainty for data intensive research facilities beyond annual cycles (P02)
- Deploy AARC Blueprint Architecture for federated authentication supporting unattended workflows (P03)

For Policy Makers and Funding Bodies:

- Recognise data-intensive research infrastructures which their long-term needs for digital infrastructure as an important stakeholder category in EuroHPC and EOSC governance (P01)
- Fund coordination mechanisms bridging thematic and horizontal infrastructure governance (P01)
- Enable strategic (multi-year) access designations aligning allocation cycles with the planning horizons for data-intensive research facilities (P02)
- Support trust framework harmonisation enabling federated authentication across European digital Infrastructures and national centres (P03)

# 4.3. Pillar 2: Architecture & Interoperability

Integrating compute, data and workflows

## 4.3.1. Challenge

European research infrastructures operate across heterogeneous computing environments that were not designed to work together. A scientific workflow may require GPU clusters for machine learning training, HPC systems for large-scale simulation, and distributed storage for multi-petabyte datasets. Currently, each resource type often requires separate interfaces, data formats, and operational procedures.

Three interconnected challenges prevent the compute and data continuum from functioning as an integrated technical capability for research:

Computing integration fragmentation: HTC and Cloud communities have established federation within their domains (WLCG federates 170 sites; EGI provides brokering across European e-infrastructures)The HPC federation is still developing, with the EuroHPC Federation Platform under development. Each HPC centre uses its own portals, schedulers, and software stacks, requiring experiments to custom-adapt to each site. Integration across the different resource types remains limited and standardised integration policies are needed to enable unified orchestration across HPC, HTC, Cloud, and emerging quantum resources (D6.1, Gaps #3, #5).

Data management fragmentation: Exabyte-scale workflows require moving data between long-term archives and transient compute resources, yet transfer mechanisms are fragmented and often rely on ad-hoc tools. No unified architectural approach exists for FAIR storage across domains, including standardised metadata and provenance tracking. Between large Tier-0/1 sites (CERN, the main WLCG and SRCNet hubs) the network backbone is generally adequate, with GÉANT and dedicated overlays (LHCONE, LHCOPN) providing high-capacity transport (D6.1 Section 4.2). Reaching beyond those core sites is the binding constraint: D6.1 Section 5.4 identifies network connectivity as a persistent technical constraint and notes that “at minimum, HPC sites need strong connectivity to WLCG centers and SRCNet nodes”, with realistic demand validation required at non-core centres. Workflow orchestration limitations: Scientific workflows increasingly span multiple facilities, yet orchestration systems remain fragmented across communities. Cross-facility workflows lack resilience and provenance capture. Resource-aware scheduling that considers data locality, compute availability, energy and cost optimisation remains immature.

## 4.3.2. Priorities Overview

Pillar 2 establishes the technical machinery enabling heterogeneous computing integration, data federation and transport, and workflow orchestration across distributed resources. Three priorities address complementary layers of the technical stack: computing abstraction (P04), data management (P05), and workflow coordination (P06).

P04: Heterogeneous Computing Integration. Provide unified access to heterogeneous computing resources (HPC, HTC, cloud, edge, and emerging quantum) so that researchers can compose workflows across providers without per-stack glue. This requires standardised integration policies adopted across the major European federations, EuroHPC Federation Platform, EGI Federation, WLCG, SRCNet, alongside national pre-exascale and exascale centres. Move beyond reliance on CLI-based batch interfaces toward API-based access mechanisms compatible with workflow engines including the LEXIS Platform (based on Apache Airflow), Common Workflow Language, Dask and HTCondor; introduce shared ontologies (with reference to Numpex) as a politically tractable first-stage requirement. Integration must connect with data infrastructures and data providers, including the European Space Agency (ESA) initiative serving Earth observation and astronomy communities, since heterogeneous compute is rarely useful without proximate data. (Detail: Annex A, P04)

P05: Data Management and Transport. .Establish cross-domain standards for federated FAIR data management, automated large-scale data movement with advance-planning and event-driven capabilities, and flexible architectures for sensitive data handling (a transversal need across data-intensive science beyond HEP/RA). For datasets too large to move repeatedly, remote computational access (where supported by the target site) complements bulk transfer rather than substituting for it, since many HPC sites block outbound connections from compute nodes. Convergence on common transfer protocols supporting both POSIX-like and object-storage access patterns, deployment of proven movement services at HPC centres, and operational observability of throughput, latency, error rates, and end-to-end staging time complete the operational picture. (Detail: Annex A, P05)

P06: Workflow Orchestration and Management. Evolve workflow orchestration into a cross-facility capability with built-in resilience, provenance capture, and support for interactive compute use cases, mixed workloads combining embarrassingly parallel tasks with tightly coupled AI processes, and unattended execution of long-running workflows. Build on D6.1’s Orchestration and Workflows capability (Workflow Management, Resource Orchestration, Task Scheduling, Provenance) and on the European workflow-management projects catalogued in the Technical Blueprint already advancing this space. Cross-priority dependencies on P01–P03 (governance, allocation, identity) carry the orchestration prerequisites identified in the D6.1 section 4.4 cross-references. (Detail: Annex A, P06)



Interdependencies: The three priorities form a layered structure. Workflow orchestration (P06) depends on heterogeneous computing integration (P04) to execute across diverse resources and on data management (P05) to stage data appropriately. Effective data transport requires understanding of compute and storage resource availability; efficient computing requires data locality awareness. Existing projects demonstrate this integration: interTwin [111] delivers workflow orchestration spanning HPC and HTC resources, whilst Rucio (the WLCG-originated distributed data-management framework, now adopted across multiple experiments) provides data management across the WLCG and SRCNet federation.

## 4.3.3. Recommendations

For Scientific Communities and Research Infrastructures:

- Down-select to a small set of supported interfaces, data-management tools, and workflow engines so that e-infrastructures can build against a stable target (P04, P05, P06)
- Ensure the capability to execute on diverse compute platforms via code porting and/or the use of high level programming frameworks (P04)
- Contribute to use cases that drive data management and transport requirements (P05)
- Engage in co-design of workflow orchestration capabilities through community working groups and with service providers (P06)

For e-Infrastructures and Service Providers:

- Implement standardised interfaces for heterogeneous resource description, discovery, and access (P04)
- Deploy interoperable data management services with FAIR-compliant metadata (P05)
- Support cross-facility workflow execution through common orchestration frameworks (P06)

For Policy Makers and Funding Bodies:

- Fund architecture-ecosystem development across European digital infrastructures (EuroHPC, EOSC, EGI) (P01, P04)
- Fund uptake of federated identity and access management for workflows and the compute, data, and network components those workflows interface with (P03, P04)
- Fund development of integration and deployment testbeds enabling validation across HPC/HTC/Cloud resources (P04)
- Support (smart) network capacity investments enabling exabyte-scale data movement (P05)
- Require interoperability standards in infrastructure procurement (P04, P05, P06)

# 4.4. Pillar 3: Software & Science Enablers

Enabling AI, portability and reproducibility

## 4.4.1. Challenge

Scientific software represents decades of accumulated knowledge and investment, yet faces growing tensions between longevity requirements and rapid technological change. Major experiments like HL-LHC and SKA will operate for 10-50 years, requiring software that remains functional across multiple generations of hardware and computing paradigms. Simultaneously, AI is transforming scientific workflows — encompassing ML/DL methods used for surrogate modelling and event classification, and increasingly LLM-based and agentic systems that can support code generation, infrastructure operations, and documentation — but most applications remain at proof-of-concept stage rather than production deployment.

Four interconnected challenges constrain the software layer bridging infrastructure and science:

- AI maturity gap: Machine learning methods are increasingly central to simulation, reconstruction, and analysis, yet most infrastructures treat them as add-ons rather than integral capabilities. Scaling from demonstration to operational deployment is challenging, often dependent on small teams rather than sustainable capacity. The gap between AI potential and AI reality represents a strategic bottleneck.
- Portability limitations: Scientific codes optimised for one architecture often cannot exploit another without substantial rewriting. The computing landscape now spans CPUs, GPUs, FPGAs and emerging accelerators, each with different programming models and each with typical evolution cycles of 2-5 years. Taking advantage of heterogeneous resources requires software that is portable across centres and automatically exploits available hardware capabilities; this can be realized by careful porting to all the architectures of interest (e.g., EESSI), or by using high-level frameworks designed for portability (e.g. Kokkos, SYCL/oneAPI, OpenMP target offload, or ALPACA).
- Reproducibility requirements: Scientific insights must be reproducible over a long period requiring preservation of complete computational context including software, workflows, documentation and execution environments. Current practices often fail to capture sufficient information for later reproduction.
- Long-term preservation: Instrument lifetimes of 10-50 years require software (together with the complete execution environment) that survives across hardware generations and organisational changes. Code developed today must remain executable and maintainable through multiple technology transitions.

### 4.4.2. Priorities Overview

Pillar 3 enables AI and accelerated computing methods, software portability across heterogeneous architectures, reproducibility of scientific insights, and code preservation across instrument lifetimes. Four priorities address the software lifecycle from development through long-term preservation.

P07: AI. Establish AI as first-class infrastructure across the European compute and data continuum. AI surrogate models can deliver two to three orders of magnitude computational cost reduction for high-impact HEP and RA use cases [10, 24]. Realising this depends on an integrated AI ecosystem that spans HPC, HTC, cloud, and instrument-edge resources, supports federated learning and FAIR model sharing, and is not concentrated in a single class of facility. The strategic lever is active engagement with European AI policy (RAISE, the AI Office, the AI Act), supported by resource-accounting reform that recognises GPU and AI-Factory allocations as pledged for the major experiments (Detail: Annex A, P07).

P08: Software Portability. Develop scientific codes using portable programming models that execute efficiently across diverse architectures (CPUs, GPUs, FPGAs, and emerging accelerators). Adopt hardware abstraction layers such as Kokkos, SYCL, and Alpaka for architecture-agnostic code; package applications in containers (e.g. Apptainer) or modules (e.g., Lmod); and follow FAIR principles for performance portability and reusable software (e.g., FAIR4RS). Build on European tooling already in production, including EESSI through the EuroHPC Centre-of-Excellence MultiXscale and the EuroHPC Federation Platform's Federated Software Catalogue, and on the EOSC Cloud Container Platform for containerised application deployment. (Detail: Annex A, P08)

P09: Open Science and Reproducibility. Ensure scientific workflows and results remain reproducible over 5-10 year timescales through preservation of (a sufficiently) complete computational context, including standardised metadata, provenance tracking, and FAIR-compliant practices for software and workflows alongside data. Connect with e.g., EVERSE for research software best-practice consolidation and with the DORA/sfdora research assessment principles for career recognition based on FAIR and Open Science contributions. Open data alliances (IVOA, IHDEA, IPDA) provide established community standards that the priority builds on. (Detail: Annex A, P09)

P10: Software Preservation and Sustainability. Ensure scientific codes remain functional across the 20-50 year lifetimes of major instruments (HL-LHC programme to its successors, SKA1 to SKA2, LOFAR through LOFAR 2.0) through systematic preservation and modernisation, with documentation as an integral part of preservation policies and Software Management Plans. HEP's long collaborative legacy from LHC informs emerging RA infrastructures such as SRCNet (still in build-out, so the comparison must be drawn carefully). The priority can be supported by e.g., EVERSE, the SPECTRUM Knowledge Hub, the ESCAPE ESFRI Cluster, and networks such as Radionet, with advocacy from DORA and the ADORE declaration of the Research Software Alliance. Promote open-source and open-standards policies to prevent sovereignty issues and avoid proprietary vendor lock-in. (Detail: Annex A, P10)

Interdependencies: The four priorities form a coherent approach to sustainable scientific software. Portability (P08) enables reproducibility (P09) by decoupling software from specific environments. AI/ML integration (P07) depends on portable software methods (P08) to deploy models across heterogeneous resources. Long-term preservation (P10) requires the documentation and standardisation practices developed for reproducibility (P09).

### 4.4.3. Recommendations

For Scientific Communities and Research Infrastructures:

- Invest in transitioning AI applications from proof-of-concept to production deployment (P07), including evaluation of LLM-assisted ("vibe-coding") workflows as a possible accelerator for porting scientific codes to new architectures and for sustaining them across hardware generations (P10)
- Engage with the RAISE AI for Science pillar to secure HEP and RA representation and contribute use cases for AI Factory pilot workloads (P07)
- Adopt portable programming models and (micro-) architecture optimized distribution for new software developments (P08)
- Implement FAIR Principles for research softwares (e.g., FAIR4RS) guidelines for software documentation and preservation  (P09, P10)
- Engage with the i.a., the EVERSE software-quality framework and the EOSC science-cluster software repositories such as the ESCAPE Open-source Scientific Software and Service Repository (OSSR) (P08, P09, P10)

For e-Infrastructures and Service Providers:

- Deploy AI platforms as integral infrastructure capabilities, not add-ons (P07)
- Provide federated-learning infrastructure for privacy- and sovereignty-constrained training, and operate AI Factory pilots that integrate HEP and RA workloads with EuroHPC AI infrastructure (P07)
- Support common software repositories and distribution systems (P08)
- Align service offerings with the community-driven and maintained  software-excellence frameworks and science-cluster software repositories (e.g., EVERSE, ESCAPE OSSR (P08, P09))
- Provide execution environments and support workflow tooling that enable reproducibility and provenance capture (P09)

For Policy Makers and Funding Bodies:



- Fund transition of AI applications from proof-of-concept to production deployment (P07)
- Fund inference infrastructure integration across distributed computing resources (P07)
- Require reproducibility standards in funded research (P09)
- Fund software preservation as an essential investment component of research and digital infrastructures (P10)
- Reform resource-accounting frameworks so that GPU and AI-Factory allocations are accepted as pledged resources within WLCG and SRCNet accounting (P07)
- Engage the AI Office on AI Act application to scientific AI, including model documentation, dataset transparency, and risk classification for safety-relevant operational AI such as detector control and observatory operations (P07)
- Sustain pan-European software-excellence frameworks such as EVERSE, science-cluster software repositories such as the ESCAPE Open-source Scientific Software and Service Repository (OSSR), and optimized software compilation, packaging and distribution such as EESSI as cross-priority enablers for software quality, training, and FAIR-for-software practices (P08, P09, P10)

## 4.5. Pillar 4: Human Capital & Responsibility

Strengthening communities, improving environmental sustainability and workforce

### 4.5.1. Challenge

The European compute and data continuum depends on people as much as technology. Technical capabilities require skilled professionals to implement, operate, and evolve them. Environmental sustainability requires not just efficient hardware, human expertise to optimise software and operations but also the management of environmental impacts over the whole life cycle of infrastructures or scientific projects, as well as the means to ensure actors at all levels (end users, data center operators, policy makers and funders) adopt environmentally responsible behaviours. Cross-infrastructure coordination requires communities that collaborate effectively across organisational boundaries.

Three interconnected challenges constrain human capital and organisational capacity:

- Workforce skills gap: The drivers described in Section 3 require expertise that is in short supply. Heterogeneous computing demands specialists in portability and performance engineering. AI/ML adoption depends on machine learning engineers who can bridge proof-of-concept and production. Sustainability goals require transverse hardware-software expertise to minimise impacts through continuous code optimisation,  impact-aware scheduling and infrastructure operation. Sustained progress requires a triangular collaboration model — researchers driving concept and validation, code developers responsible for implementation and optimisation, and digital-infrastructure specialists owning operations and performance — funded as a multi-year commitment rather than as ad-hoc project additions. Community surveys indicate that most research groups identify technical capacity as a concern, yet lack resources and expertise to address it.
- Career structure misalignment: Scientific career structures do not recognise or reward contributions to infrastructure, software development, or operational efficiency. Research software engineers, who bridge domain science and computing expertise, typically lack defined career paths. This creates retention challenges as the same skills command premium salaries in the commercial sector. The Amsterdam Declaration on Funding Research Software Sustainability identifies recognition of software as a first-class research output, but operational uptake has been uneven: most large-scale digital infrastructures provide only limited code-optimisation support, and deep domain-level RSE work is rarely embedded in their core remit. Partial implementations exist via EuroHPC Centres of Excellence, RI Train, and HPC Train, but no European-level funding line yet sustains the triangular researcher–RSE–DI collaboration the priority calls for. The Declaration treats recognition as a research contribution as essential to workforce sustainability.
- Fragmented community coordination: Research infrastructures and the federated digital infrastructures serving them are coordinated unevenly: WLCG and the Grid stack have a mature multi-party governance, but the coupling between data-intensive science needs and EuroHPC remains under-articulated, and structured engagement between scientific RIs and the broader European digital-infrastructure providers is still partial. Scientific requirements are often invisible when infrastructure decisions are made. Effective co-design requires sustained collaboration that current funding models do not reliably support.

### 4.5.2. Priorities Overview

Pillar 4 builds collaborative communities through co-design processes, establishes environmental sustainability through impact-aware management of facilities, and develops career profiles combining domain science with computing expertise. Three priorities address the human and organisational dimensions of the continuum.

P11: Community Collaboration and Co-Design. Establish structured mechanisms for scientific communities to engage with infrastructure development, building on the SPECTRUM CoP and existing networks (APPEC, ASTRONET, RadioNet, JENA, ESCAPE). Strengthen the still-not-fullyarticulated alignment between EOSC (as the European Data Space for research) and EuroHPC through structured dialogue, coordination, and mutual visibility, leveraging the AI Factories' explicit obligation under the AIF Council Regulation to support the official European Data Spaces. Here there is an important role for the European Commission and its DGs (DG CONNECT and DG RTD) to ensure programmatic alignment. The priority addresses human-capacity needs alongside coordination - stable roles, recognised responsibilities, organisational incentives, and retention of key engineering and software expertise - since organisational, political, and workforce factors are often as or more decisive as the underlying technology. (Detail: Annex A, P11)

P12: Environmental Sustainability. Embed environmental accountability for multiple impacts into research computing operations, using lifecycle assessment methodologies (ISO 14040:2006 and ISO 14044:2006 on Environmental Management, Life Cycle Assessment; the EU Product Environmental Footprint (PEF) method) to track impacts across procurement, fabrication, operation, and end-of-life. For this accounting to drive behaviour rather than displacement, it must be paired with policy on research conduct (procurement criteria, allocation rules), otherwise raised EU-tracked costs risk pushing workloads to non-European providers and net-out the sustainability gain.

Allocate environmental impacts as credits flowing from policy makers to data centres and end users, replacing core hours and abstract metrics with visible impact accounting. The credit metaphor, rather than direct euros, allows top-down political targets (e.g. EU Green Deal limits) to flow through the allocation chain without entangling with market-cost dynamics; the two metrics are complementary, not substitutable, that supports responsible behaviour.

Multi-objective optimisation balancing performance with carbon emissions, energy consumption, and resource utilisation is a candidate output, with best-practices for European data centres (best practices and shared reference patterns rather than rigid blueprints, given the diversity of facility designs) an option for community discussion. The priority addresses growing requirements for environmental responsibility as a condition of public funding and social licence to operate. (Detail: Annex A, P12).

Relaxing the environmental tension requires addressing the whole value chain leading to the generation of science data products in order to reduce the environmental impact: to jointly optimise the research problem (recognising that responsibility for resource-appropriateness lies with the researcher's problem framing as well as with software and hardware) in combination with software and hardware, to make maximal use of minimal resources, to tailor execution to environmental impact in the sense of choosing what resources to use where and at what moment in time, to reduce digital waste through the reuse of data and software, framing failures during execution and resilient execution by adapting to changing operating conditions. Optimising scheduling and hence avoiding idle cycles is equally important. This relies on combining tools, policies and a structured collaborative effort between research and digital infrastructures.

Key to their successful application is a change of paradigm which consists in allocating environmental impacts at all levels, instead of core hours for end users and only funds to data centers. This change of metric is intended, on one hand, to ensure these impacts are visible in order to foster responsible behaviours by the actors and, on the other hand, to establish a framework allowing a high-level allocation made by policy makers in line with European or national objectives to then be used by data centers, typically through hardware procurement and operations, with the latter itself resulting from impact allocation to end users. These allocations serve as credit which each actor is then free to manage (within a framework that is collaborative set in e.g., a Commons approach by science funders, research infrastructures and digital infrastructures) and consumed by arbitrating between their priorities. At the level of the compute and data continuum, this is expected to require coordination between institutional actors in order for the continuum to not just provide extended resources to users but also a lever to reduce the global and local environmental impacts of their activities.

P13: Skills, Careers and Training. Develop the hybrid expertise combining domain science with computing skills that the continuum requires, with training programmes aligned with emerging drivers (AI/ML integration, heterogeneous computing, sustainability practices), delivered through a mix of integrated curricula and AI-assisted training tools (which can extend reach where curriculum capacity is the bottleneck), and career frameworks following the Research Software Engineer model. The current arrangement carries an avoidable systemic cost: trained experts leave at later career stages, taking with them critical, slow-to-rebuild institutional knowledge, and forcing complete system replacements that would not otherwise be needed. The intent is not to compete with industry on salary but to retain the specific, hard-to-substitute knowledge that the long-lived experiments (HL-LHC, SKA) actually depend on. A modest number of stable, well-placed career positions would deliver a better return on the substantial investments already committed to scientific computing. Short-term actions may leverage current EU instruments (e.g., RI Train, HPC Train, EuroHPC Centres of Excellence); medium-term actions integrate training into European university curricula (Detail: Annex A, P13)

Interdependencies: Community collaboration (P11) provides the foundation for this pillar and connects across all pillars through co-design processes that shape governance decisions (Pillar 1), technical architecture (Pillar 2), and software priorities (Pillar 3). Within Pillar 4,



P11 coordinates sustainability practices (P12) and enables knowledge transfer for skills development (P13). These priorities create the organisational foundation on which technical capabilities depend.

### 4.5.3. Recommendations

For Scientific Communities and Research Infrastructures:

- Increase and incentivise participation in co-design processes and CoPe Working Groups (P11)
- Integrate and reward environmental impact awareness into software development practices (P12)
- Recognise and value research software engineering contributions within projects and collaborations (P13)

For e-Infrastructures and Service Providers:

- Maintain community engagement mechanisms through dedicated liaison roles and coordination structures (P11)
- Implement a sustainability governance through impact reporting and impact-driven resource allocation (P12), supported where possible by EU-level regulation that institutions can act on (the post-CSRD/Omnibus reporting landscape is still settling, so the regulatory hook for research computing remains to be defined)
- Provide automation (e.g, AI-driven), training and documentation enabling efficient use of heterogeneous resources (P13)

For Policy Makers and Funding Bodies:

- Fund sustained community coordination mechanisms beyond project timelines (P11)
- Define admissible impact budgets for infrastructures and evaluate investments and operation in this frame (e.g., by implementing reporting requirements at the EU-level) (P12)
- Establish career frameworks for research software engineers (across emerging skill areas including AI/ML, performance engineering, and scientific code preservation) and for data stewards (responsible data management, hierarchical-storage management, and long-term curation) (P13).

# 5. Investment Areas



Realising the compute and data continuum requires coordinated investment across five areas. This section maps the thirteen strategic priorities to those areas and to the pillar framework introduced in Section 4, identifies the European funding instruments available, and highlights investment patterns that traditional hardware-and-floor-space budgeting tends to underestimate (coordination labour, federation-software engineering, RSE effort, and long-horizon data curation).

## 5.1. Investment Areas, Pillars, and Priorities

Table 5.1 maps the five investment areas to pillars, priorities, dominant expenditure types, and the specific investments involved.

Table 5.1: Investment areas mapped to pillars, priorities, and expenditure types.

| Investment Area | Associated Pillar | Primary Priorities | Dominant Expenditure Type | Key Investments |
|---|---|---|---|---|
| **Computing Infrastructure** | Pillar 2 (Architecture and Interoperability) | P04 (primary), P07 (secondary) | Capital + operational | Integration of thematic and national interfaces (especially community-specific interfaces for data-intensive science) with the EU-level federation ecosystem (EuroHPC Federation Platform, EOSC); enabling technologies that bridge thematic edge/continuum resources with EU-level capacity; cross-facility workflow orchestration capacity at execution sites; software distribution infrastructure; large-scale AI training infrastructure and distributed inference resources |
| **Data Infrastructure** | Pillar 2 (Architecture and Interoperability) | P05 (primary), P09 (secondary) | Capital + operational | Tiered storage systems with standard access protocols; high-bandwidth transfer networks with advance reservation; FAIR repositories with standardised metadata and provenance tracking. European Data Spaces target a single market for data, but the distributed research-data infrastructure that this agenda depends on is not market-driven; a more active EC coordination role is needed to address this market failure for research data. |
| **Software and Tools** | Pillar 3 (Software and Science Enablers) | P06, P08, P10 (primary), P07, P09 (secondary) | Sustained development + operational | Hardware abstraction layers (e.g. Kokkos, SYCL), modules (e.g., Lmod) and containers (e.g. Apptainer); CI/CD, compilation and software distribution infrastructure; workflow management systems and orchestration tooling; domain repositories and modernisation processes for software preservation |
| **Governance and Coordination** | Pillar 1 (Policy, Trust and Governance); Pillar 4 (P11, P12) | P01, P02, P03, P11, P12 | Coordination | Liaison mechanisms and advisory representation; multi-year allocation frameworks; federated identity, authentication, and authorisation infrastructure (based on the AARC Blueprint Architecture, full AAI stack including attribute release and token translation, with eduGAIN as the underlying identity federation); sustained co-design processes and European/international coordination; sustainability governance and evaluation |
| **Human Capital** | Pillar 4 (Human Capital and Responsibility) | P13 | Operational (permanent positions) | Training curricula for emerging skills; career frameworks for research computing and data professionals (including research software engineers, data specialists, and domain scientists dedicated to computing); retention strategies competing with industry; adoption and incorporation of AI (e.g., agentic AI), requiring sustained development effort in addition to training and recruitment |

Cross-cutting: P09 (Open Science and Reproducibility) spans computing, data, software, and governance investments. P12 (Environmental Sustainability) applies across all infrastructure and operations. Operational expenditure is present in every priority; the table records its dominance where no other type is primary.

## 5.2. Investment Profiles Across Implementation Phases

The investment profile changes across the roadmap phases described in Section 6. Aligning funding instruments with actual expenditure patterns requires recognising this shift.

In the short-term foundation phase (years 1-3), investment is dominated by coordination costs and design work. Establishing governance forums (P01), bootstrapping trust frameworks (P03), conducting gap analyses (P04, P05, P06), and launching community engagement (P11) require sustained staff effort rather than capital procurement. The primary resource needed is people: liaison officers, policy specialists, AAI operations staff, and community coordinators. Pilot projects and testbeds require modest computing and data resources alongside engineering capacity.

In the medium-term integration phase (years 3-5), the profile shifts toward technical deployment. Prototyping federated data management across research infrastructures (P05), deploying workflow orchestration (P06), validating heterogeneous computing integration (P04), and consolidating portable software stacks (P08, P10) require both infrastructure investment and sustained engineering effort. Multi-year allocation frameworks (P02) move from design to operational deployment, requiring structural funding for accounting systems and MoU-based coordination.

In the long-term maturation phase (years 5+), the dominant need is sustained operational funding. Production-grade federated AI/ML environments (P07), software preservation repositories (P10), reproducibility services (P09), and sustainability reporting frameworks (P12) must operate reliably over multi-decade scientific programme timescales. Career frameworks for research computing and data professionals (P13) require permanent positions, not project funding. To ensure that operational funding is applied in an optimal and effective manner over time, sustained development and innovation funding are also needed but these do not dominate the overall cost.

This progression means that short-term investment can largely be addressed through project-based funding (e.g., individual Horizon Europe and FP10 calls, national pilot programmes), but medium- and long-term success depends on structural funding commitments that outlast individual project cycles and requires a longer-term programmatic approach (e.g., via EuroHPC, EOSC).

## 5.3. Mapping of Priorities Across European Funding

Investment in the compute and data continuum draws on multiple European and national funding streams. No single instrument covers all investment areas or phases; effective implementation requires coordinated use of complementary sources. Section 3.1 describes the broader European ecosystem within which these instruments operate; this section focuses on their funding mechanics, committed amounts, and alignment with investment areas.

National funding provides the majority of computing capacity for European research infrastructure. The resources underpinning WLCG, SRCNet, LOFAR, EGI, and national HPC centres are predominantly nationally funded. Without sustained national commitments, long-term operational continuity is not feasible; EU Framework Programmes typically provide approximately 10% of research infrastructure budgets.

EuroHPC Joint Undertaking has committed approximately EUR 3.08 billion across the 2021-2027 Multiannual Financial Framework (combining EU contributions with matched contributions from participating states and private partners) for digital infrastructure at an EU-level. This covers exascale supercomputing systems, AI Factories, quantum computing platforms, and the EuroHPC Federation Platform (under development). EuroHPC is directly relevant to priorities P04 (heterogeneous computing), P07 (AI/ML), and P03 (federated identity through the Federation Platform). Access for data-intensive science communities remains governed by EuroHPC's allocation mechanisms, which Priority P02 seeks to align with multi-year scientific planning horizons for research infrastructures.

The AI Continent Action Plan (2025) and the proposed Cloud and AI Development Act (CAIDA, target publication 2026 within the EU technological sovereignty package) provide the principal funding pathway for P07 (AI), with over EUR 10 billion committed to AI Factories and over EUR 20 billion for AI infrastructure overall. The RAISE programme coordinates AI capabilities for science across the EU; HEP and RA representation in the RAISE AI for Science pillar is the actionable engagement route. IPCEI-CIS provides a co-funding pathway for sovereign federated cloud relevant to scientific AI infrastructure, and the Berlin Declaration (November 2025, signed by 17 Member States) accelerates implementation of IPCEI-CIS and aligned AI infrastructure investments.

European Open Science Cloud (EOSC) covers priorities P05 (data management), P09 (open science), and P03



(federated identity). EOSC operates as a co-programmed European Partnership: EOSC-A Members have pledged approximately EUR 500 million in in-kind contributions to co-finance the partnership over 2021-2030, matching public funding allocated through the EOSC partnership envelope. The EOSC EU Node and Federation are under active development.

Horizon Europe (2021-2027) funds research infrastructure through dedicated call topics including INFRA-DEV (design and development), INFRA-TECH (technology), INFRA-SERV (services), and INFRA-EOSC (EOSC integration). These instruments support short-term design, prototyping, and pilot work. ESFRI infrastructures and ERICs benefit from explicit eligibility criteria in many calls, facilitating consortia formation. Current projects relevant to SPECTRUM priorities include ENSURE (INFRA-TECH, relevant to P02 federated allocation design), interTwin (INFRA-TECH, HPC federation, relevant to P04), ODISSEE [140] (INFRA-TECH, relevant to P04, P05, P06, P07, P08), and EVERSE (INFRA-EOSC, research software quality, relevant to P08, P09, P10).

The Horizon Europe 2026-2027 Work Programme for Research Infrastructures includes structural developments relevant to this agenda. The INFRA-DEV programme now supports thematic cluster consolidation across six ESFRI domains, with "Data, Computing and Digital Research Infrastructures" introduced as a domain in its own right. This creates a formal home within the ESFRI cluster framework for the kind of federated computing and data infrastructure SPECTRUM addresses. Alongside INFRA-DEV, the INFRA-TECH programme is the principal vehicle for sustained technical development of distributed research-IT, and the INFRA-EOSC line plus the EuroHPC JU work programmes are the two main EU-level streams whose direction the data-intensive science communities need to influence.

The INFRA-SERV programme groups ESFRI domains into integrated access areas, combining "Physical Sciences and Engineering" with the data and computing domain within a single area. We note that there is an absence of INFRA-SERV projects. This appears to be due to (at least) two reasons: (a) providing services for a short duration for a widening group is not aligned with service provisioning and investment cycle for research and digital infrastructures, (b) CAPEX (depreciation) costs are typically not eligible and hence access would be provided a net-loss per unit via e.g., virtual and/or transnational access.

A project that clusters the shared digital-infrastructure needs of HEP and RA research infrastructures, federated identity, data movement, workflow orchestration, sustainability accounting, and channels them into the EuroHPC, EOSC, and EGI roadmaps would fit this funding structure directly, without implying any merger or deep integration at the RI level. These programme directions provide a funding pathway for the short-term foundation activities described in Section 6.

Framework Programme 10 (FP10, 2028-2034) is under negotiation. The proposed structure includes four pillars, with research and technology infrastructures positioned in a new Pillar IV (European Research Area) with an estimated allocation of approximately EUR 16 billion, of which research and technology infrastructures may receive approximately EUR 8 billion. The proposed FP10 "20% construction" allocation refers to new RESEARCH infrastructures (ESFRI Roadmap entries), not to digital research infrastructures — EuroHPC continues to operate at its existing 50% EC / 50% Member State co-funding model. The European Competitiveness Fund will be linked to FP10 to ensure that European industry benefits from research results. The European Parliament and the Council of the EU (in the Warsaw Declaration) have expressed that R&I must be funded by a specific instrument. A new provision allows the European Commission to finance up to 20% of construction costs for critical new infrastructures, which could apply to new computing and data services.

ERA Act and Policy Agenda. The forthcoming European Research Area Act will establish a binding regulatory framework for cross-border research cooperation, realising the "fifth freedom" of free movement of researchers, knowledge, and technology. The ERA Policy Agenda (2025-2027) includes active actions that map to SPECTRUM priorities: research infrastructures (25 Member States participating), research careers (27 MS), AI in science (23 MS), enabling open science (24 MS), and research security (26 MS). The ERA action on Research Management as a distinct career path is particularly relevant to Priority P13.

ESFRI provides strategic direction and legitimacy through its Roadmap process. While ESFRI has no direct funding authority, ESFRI status facilitates discussions with national ministries and provides eligibility advantages in EU calls. The ESFRI Report on Funding of Research Infrastructures (2025) identifies persistent challenges including long-term operational sustainability, coordination between funding sources, and human resources. Its recommendations align closely with SPECTRUM priorities P01 (governance), P02 (allocation), and P13 (workforce).

European Regional Development Fund (ERDF) can contribute to research infrastructure construction, with co-funding rates up to 85% in some regions. Operations funding through ERDF is more constrained.

Beyond these instruments, the EC's push towards integration across Research, Technology, and Digital Infrastructure categories (Section 3.1) may open additional funding pathways for shared digital capabilities serving science, particularly where collaboration between RIs and TIs is encouraged.

## 5.4. Underestimated Investment Needs

Three categories of investment are consistently underestimated in research and digital infrastructure budgets. In particular these categories require structural funds and without them, capital investments fail to deliver returns Coordination costs. Every pillar of this agenda requires coordination investment: liaison capacity bridging governance structures (P01), policy specialists aligning legal and privacy constraints (P03), community coordinators sustaining co-design processes (P11), and sustainability governance staff (P12). This coordination can be delivered through in-kind contributions from participating organisations or through existing institutional frameworks, but it must be sustained beyond individual project cycles (typically 3-year). The ESFRI Funding Report (ESFRI LTS-WG, 2025) identifies discontinuity in coordination funding as a persistent challenge.

Software maintenance and support. Software makes infrastructure usable. The Amsterdam Declaration on Funding Research Software Sustainability calls for recognition of software maintenance as ongoing infrastructure cost rather than a one-time project deliverable. Priorities P06 (workflow orchestration), P08 (portability), P09 (reproducibility), and P10 (preservation) all depend on sustained software development and maintenance that current project-based funding models do not reliably support. The shift from "software as a project output" to "software as a maintained service" is a funding model change, not a technical one.

Human capital retention. Research infrastructures cannot compete with industry on salary for the technical skills this agenda requires: heterogeneous computing specialists, machine learning engineers, security architects, sustainability analysts. Priority P13 addresses career frameworks and training, but the underlying investment need is for permanent positions with competitive conditions. Training programmes create belonging and community, but retention requires institutional commitment. The ERA Policy Agenda's emerging action on Research Management as a career path signals policy-level recognition; similar efforts are needed for research software engineers and infrastructure technical staff Capital expenditure on hardware yields returns only when coordination, software, and people investments are sustained alongside it.



# 6. Multi-Annual Roadmap



This section sequences the thirteen priorities across time horizons, reflecting dependencies between priorities, realistic implementation timescales, and alignment with European funding cycles. The roadmap spans 2026-2036 and beyond, covering the preparation and operational phases of HL-LHC and SKA Observatory.

## 6.1. Phasing Rationale

Priority sequencing follows four principles:

- Governance and trust before technical integration: Cross-infrastructure coordination (P01), federated identity (P03), and allocation frameworks (P02) must be in place before computing, data, and workflow integration can proceed at scale. These governance foundations enable interoperability at legal, organisational, semantic, and technical layers.
- Infrastructure before application: Computing and data capabilities (Pillar 2) and the software that exploits them at scale (Pillar 3) develop in lockstep through a phased, agile, validated scale-up; infrastructure capacity needs to be at least at pilot maturity before scaled-out application deployment becomes useful.
- Sustained investment: Human capital (P13), community engagement (P11), and environmental sustainability (P12) require continuous attention across all phases. These are enabling conditions, not sequential steps.
- Alignment with European funding cycles: The transition from Horizon Europe (ending 2027) to FP10 (2028-2034) creates a natural boundary. Short-term actions should use remaining Horizon Europe instruments for design and piloting. Medium-term actions should target FP10 calls and the forthcoming European Competitiveness Fund for deployment and integration. Long-term actions require structural commitments from national programmes and multi-year (e.g., MoU-based) frameworks that outlast individual funding cycles

## 6.2. Roadmap Overview

Figure 6.1 shows all thirteen priorities across the full timeline, organised by pillar, alongside European funding cycles and scientific programme timescales. Filled cells indicate where each priority's primary effort falls; outlined cells show phases where the priority is active but secondary. Pillar 4 priorities span all phases continuously.

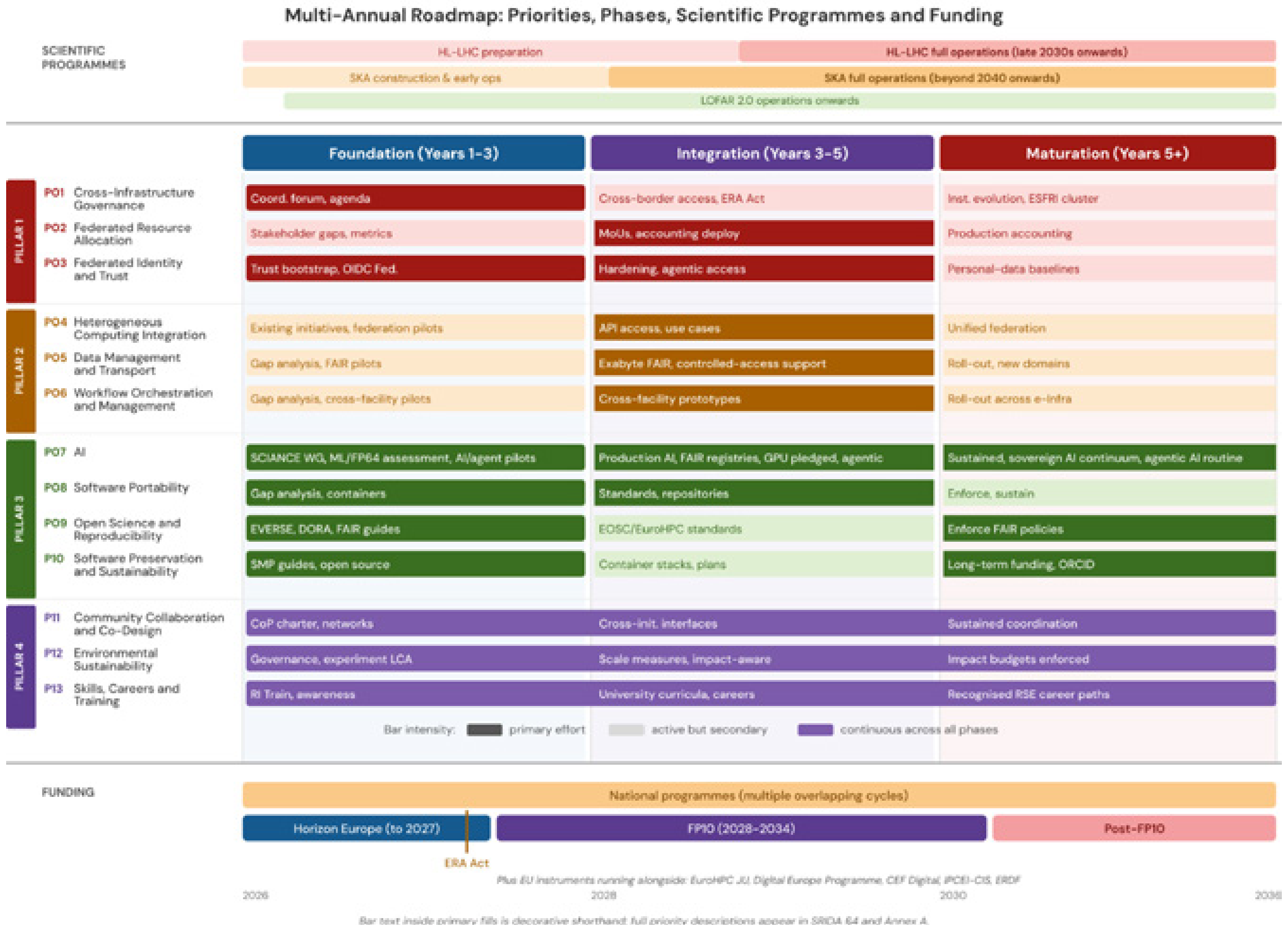



Figure 6.1: Multi-annual roadmap.

## 6.3. Detailed Phase Descriptions

### 6.3.1 Phase 1: Foundation (Years 1–3)

The foundation phase establishes the institutional and trust prerequisites without which technical integration cannot proceed at scale. Investment in this phase is predominantly coordination, development and people, not capital. The Horizon Europe 2026-2027 Work Programme provides a concrete funding pathway for foundation-phase actions, including thematic-cluster coordination calls and large-scale integrated-access calls covering the new "Data, Computing and Digital Research Infrastructures" domain — with the caveat that integrated-access call cost-eligibility currently excludes capital expenditure, so the pathway is workable for coordination, software and trans-national access but not for hardware investment. Computing and Digital Research Infrastructures" ESFRI domain together with physics-cluster access. Call identifiers are listed in Section 5.3 and the accompanying policy brief.

Governance and coordination (P01). Establish a coordination forum bringing together thematic and multidisciplinary digital infrastructures with policy maker representation. Define its substantive agenda: procurement alignment, innovation coordination for data-intensive science, and bridging the operational-versus-project funding mismatch. Map existing coordination mechanisms and identify gaps the forum addresses. Develop shared metrics for infrastructure utilisation and cross-infrastructure workflow performance.

Federated identity and trust (P03). Bootstrap trust governance across participating infrastructures. Agree operational KPIs for federated authentication. Review site-local policy and risk-acceptance barriers to agentic access (service accounts), which currently prevent automated workflow execution across HPC centres. Adopt the AARC Blueprint Architecture as the architectural baseline and evaluate OIDC Federation as the scalable technical mechanism for OIDC and OAuth deployments.

Resource allocation design (P02). Continue stakeholder conversations to close the two identified gaps: long-term assured access and planning of resource allocations, and e-Infrastructure-wide accounting valid across the whole infrastructure landscape. Establish the interoperability framework for resource accounting: agree metrics and checkpoints, identify tasks and responsibilities, deploy first testbeds and feasibility studies. Identify appropriate Horizon Europe calls for design and prototyping of federated accounting systems.

Computing integration (P04). Build on existing initiatives (e.g., WLCG federation, EGI HPC integration, EuroHPC Federation Platform). Agree on a common descriptive language, shared ontologies mapping across domain-specific vocabularies and HPC/HTC systems, as a first-stage requirement for resource sharing. Begin policy discussions on standardised interfaces across HPC, HTC, cloud, and emerging quantum providers, and on extending coverage to public scientific data archives of broader European interest, since heterogeneous compute is rarely useful without proximate data. Move beyond reliance on a single CLI-type batch interface and towards API-based access mechanisms compatible with workflow engines.

Data and workflow pilots (P05, P06). Conduct gap analyses for selected use cases. Identify the interfaces between research infrastructures, e-Infrastructures, and community APIs that constrain workflow execution. Pilot representative use cases for FAIR-compliant cross-RI data federation, automated data staging from archives into HPC environments, GDPR-sensitive workflows using secure enclaves or logically separated data zones, and cross-facility workflow orchestration . Begin developing operational observability across federated digital services - throughput, latency, error rates, and end-to-end staging time.

AI/ML pilot (P07). Pilot validation of core AI/ML capabilities — AI/ML platforms, federated storage, unified authentication, module-based and containerised software stacks, and workflow orchestration — across selected HPC and cloud sites. Engage early with EuroHPC AI Factories and the European Commission's AI policy programmes (including RAISE) so HEP and RA communities are represented under the AI for Science pillar, and address the structural gap that GPU resources are not yet accepted as 'pledged' allocations for LHC experiments.

Software foundations (P08, P10). Complete a gap analysis of existing portability tools and systems. Establish and document common software stacks for major research infrastructures. Increase adoption of containerisation for software distribution. Develop guidelines, best practices, and policies for research software preservation as an integral part of Software Management Plans, drawing on the Practical Guide to SMPs as a template. Promote open-source and open-standards policies to prevent sovereignty issues and avoid proprietary vendor lock-in.

Open science and reproducibility (P09). Consolidate guidelines and best practices for FAIR reproducible workflows under Open Science, in connection with EVERSE and aligned with the DORA/sfdora research assessment principles. Implement cross-domain collaboration tools (execution environments, harmonised metadata standards, workflow systems) that streamline findability, analysis, reproducibility, and provenance across European digital infrastructures. Establish Open Science-based career-recognition best practices and policies.

Community and workforce (P11, P13). Strengthen the SPECTRUM CoP by developing its charter, ensuring continuity, and aligning with existing networks (APPEC, ASTRONET, RadioNet, JENA, ESCAPE). Develop shared metrics, tools, and methodologies for usage, performance, and efficiency assessment, and create a shared knowledge base. Start delivering actions through current and future EU instruments . Raise awareness across academic and research institutions of the need for stable career paths for research software engineers and computing experts, framed not as competing with industry salaries but as recovering the value lost when expertise leaks at later career stages.

Environmental sustainability (P12). Institute sustainability governance at data-centre and governing-body levels through dedicated committees. Put available solutions into practice to reduce IT environmental impacts and provide some relief to the environmental crises in the short term, with limited disruption. In parallel, plan, develop, and experiment with the solutions required for the change of paradigm: lifecycle assessment tooling (cradle-to-grave), forecasting impacts, and impact-aware scheduling that exposes environmental costs to users at job-submission time alongside core hours.

**End of Phase 1**
Governance forum operational; AARC Blueprint Architecture adopted as baseline and OIDC Federation evaluation completed for federated authentication, with first cross-infrastructure authentication pilots running between participating sites; FAIR workflow and software preservation guidelines established.

### 6.3.2. Phase 2: Integration (Years 3–5)

With governance and trust frameworks in place, investment shifts towards scaling up the infrastructure deployment and its use, standards consolidation, and operational testing.

Multi-year allocation (P02). Move from testbed to production: pilot e-Infrastructure-wide resource allocations and quotas. Deploy the standardised accounting system progressively across resource providers, infrastructures, and research communities. Establish Memoranda of Understanding for multi-year access aligned with the multi-decade timescales of HL-LHC, SKA Observatory, and other ESFRI scientific programmes.

Heterogeneous computing (P04). Deliver first API-based access implementations across initiatives, addressing the different requirements of HTC batch processing, AI training, inference, and interactive analysis. These APIs are designed to consume the federated identity and authorisation tokens delivered by P03 in this phase (see Phase 2 P03 actions), so that API access and federated AAI converge rather than diverge. Iterate based on lessons learned from foundation-phase pilots. Begin coordinating with public scientific data archives of broader European interest so that heterogeneous compute integrates with proximate data services rather than being delivered in isolation.

Data management (P05). Prototype and demonstrate integrated FAIR data management across multiple research infrastructures and e-Infrastructure providers. Validate architectures for exabyte-scale data volumes. Expand to high-frequency data ingestion (radio astronomy streaming, gravitational-wave low-latency data, WLCG bulk LHC data ingestion to Tier-1s), multi-protocol storage access, and, where applicable for the small share of workflows touching GDPR-sensitive data, secure enclaves or logically separated data zones. Develop standards for FAIR metadata schemas, access-control and identity integration, and operational interfaces for cross-domain data movement. Consolidate the Phase 1 draft AARC Blueprint Architecture (AARC BPA — see also D6.1 §3 for the technical blueprint context) for European data management and transfer services into a community-reviewed version.

Workflow orchestration (P06). Prototype and demonstrate cross-facility orchestration with selected e-Infrastructure providers. Validate fitness-for-purpose across multiple workflow classes (HEP, SRCNet, LOFAR, and others). Explore standardisation of orchestration interfaces and converge on a limited set of interoperable workflow systems with support for unattended execution, provenance tracking, and cross-infrastructure scheduling. Consolidate the Phase 1 draft Blueprint Architecture for workflow systems into a community-reviewed version, scoped to where convergence is feasible across HEP, RA, and adjacent communities.

Software portability and preservation (P08, P10). Ensure adequate tooling is available to support hardware abstractions at all levels. Document commonly-accepted standards for portability across the (SPECTRUM) Compute and Data Continuum. Publish software stacks through common repositories. Consolidate standardised, community-based module stacks and container images across next-generation heterogeneous hardware. Develop plans for software sustainability as an integral part of scientific programme preparation, particularly for major experiments. Integrate research-software training into European university curricula and evaluate AI support.

Federated identity hardening (P03). Harden the trust framework established in Phase 1: onboarding and offboarding procedures, audit expectations across infrastructures, and operational Service Level Agreement (SLAs) for credential issuance and revocation. Enable agentic access (service accounts for unattended workflows) and community-coordinated authorisation, including any compensatory controls or trust-framework arrangements required by participating sites. Federated authentication, established in Phase 1 pilots, moves to widespread production deployment in this phase, so that the API-based access in P04, the data-federation services in P05, the workflow orchestrators in P06, and the AI/ML platforms in P07 all consume the same identity layer. Establish a trust and security coordination framework that includes federated community participants, with continuous readiness evaluation.

Open science (P09). Consolidate aligned interoperability standards between major EU initiatives (EOSC, EuroHPC JU, ESFRI). Consolidate FAIR-by-design services available across borders and across scientific disciplines. Develop preservation plans for data, software, and results as an integral part of scientific research preparation. Promote Open Science as part of educational curricula in university programmes.

AI/ML integration (P07). Pilot validation of core capabilities: AI/ML platforms, federated storage, unified authentication, containerised software stacks, and workflow orchestration across selected HPC and cloud sites. Integrate cloud and HPC access models. Engage with the European Commission's AI policy programmes (including RAISE) to ensure HEP and RA communities are represented under the AI for Science pillar.

Governance consolidation (P01). Consolidate coordination to deliver cross-border access to data and compute services, piloting integrated access across thematic and digital infrastructure clusters. Develop policy recommendations on the operational-versus-project funding mismatch, aligned with the ERA Act framework. Assess options for institutional evolution, including formal RI representation within EuroHPC and EOSC governance and coordination through the emerging ESFRI digital cluster. Establish joint working groups on cross-cutting technical challenges connecting the forum to Pillar 2 implementation. Community coordination (P11). Foster interaction, coordination, and potential convergence with related domain initiatives. Formalise cross-initiative interfaces: maintained inventory of European and international programmes, infrastructures, and policy processes; named liaison functions; periodic review points.

Environmental sustainability (P12). Consolidate the experimental measures from Phase 1 and put them into wide-scale practice across European data centres and research infrastructures. Scale impact-aware scheduling so it is exposed to users at job-submission time alongside core hours. Mature the lifecycle assessment tooling so that procurement, maintenance, and usage impacts are tracked consistently and can be compared across sites. Begin experimenting with the more complex measures: environmental impact budgets allocated by policy makers, and credit-based allocation flowing to data centres and end users. Consolidate sustainability metrics (Power Usage Effectiveness -PUE-, Carbon Usage Effectiveness -CUE-, embedded carbon, water usage, repairability, HEPscore/watt) to enable their usage as (standard) procurement criteria.

Workforce (P13). Continue awareness campaigns. Work with academic institutions and research organisations to establish dedicated computing career paths through future project proposals and integration into educational university programmes and curricula. Establish shared training catalogues and formal recognition of training across communities. Facilitate contacts between research computing professionals and industry through secondment opportunities. Evaluate the use of AI to support the existing workforce and help train/support the future workforce.

**End of Phase 2**
Scientific workflows span HPC, HTC, and cloud through unified interfaces with automated data staging; MoUs are operational for supporting multi-year allocations.

### 6.3.3. Phase 3: Maturation (Years 5+)

By year five (approximately 2032), the integrated continuum should reach production-scale operation, just before HL-LHC enters full luminosity from the late 2030s and as SKAO advances through its phased construction toward full operations beyond 2040. Phase 3 is therefore a phase of maturation, consolidation, and adaptation: hardware, software, and AI capabilities will continue to evolve substantially over the next decade, and the priorities below define the directions to maintain rather than the endpoints reached.

AI/ML at scale (P07). Operate sustained, interoperable European AI/ML environments with production-grade federated AI and AI-driven services, automated provenance tracking, formalised co-design and governance models, and embedded AI/ML as a strategic capability supporting Europe's exascale science leadership. HEP and RA communities have ongoing representation in EU AI policy programmes including RAISE Reproducibility and open science (P09). Enforce FAIR standard policies for publicly funded research. Provide long-term funding for data and software preservation. Establish Open Science-based career recognition as standard practice across European research institutions, supported by DORA-aligned research assessment frameworks.

Software preservation (P10). Ensure long-term funding for software preservation across scientific disciplines. Define career paths for research software engineers, with contributions made visible and assessable through persistent identifiers and citation practices (ORCID for people, DOIs for software releases via Zenodo or similar, CodeMeta and CITATION.cff for software metadata) and through DORA-aligned research assessment frameworks adopted by funders and institutions. Software stacks remain executable across hardware generations through sustained maintenance, with documentation embedded as a first-class element of preservation policies and Software Management Plans. Invoke AI support for the software process from design to operations, preservation and re-use.

Federated allocation in production (P02). Federated, flexible resource allocation processes are widely adopted.



Standardised resource accounting is routinely used in production across European e-Infrastructures, and funding bodies receive coordinated evidence on the scientific return from their investments.

Computing federation (P04). Unified federation across HPC, HTC, cloud, and emerging quantum platforms, based on lessons learned from earlier phases. Standardised interfaces and shared ontologies enable scientific workflows to exploit the full compute continuum with minimal per-site adaptation, confined to performance tuning rather than re-engineering.

Data management and transfer (P05). Roll out federated data management and automated transport across European e-Infrastructures. Onboard additional scientific communities beyond HEP and RA, including new scientific domains. Ensure sustainability through collective governance, coordinated investment, and community-driven standards (e.g., using a digital commons approach).

Workflow orchestration (P06). Roll out cross-facility workflow orchestration across European e-Infrastructures, expanding the community base and supporting interactive use cases at scale.

Software portability (P08). Require that all software intended for reuse beyond its original authoring group, that is, libraries, frameworks, workflow components, and tools that support scientific publications, is published through established standards and repositories. Ensure sustained long-term support for enabling technologies and portability frameworks (e.g., Kokkos, SYCL, Alpaka, EESSI, CernVM-FS, Apptainer).

Governance maturation (P01). Assess the impact of governance coordination. Determine whether informal coordination or a formalised institutional structure provides the right balance for multi-decade planning, drawing on options including formal RI representation within EuroHPC and EOSC governance and coordination through the ESFRI digital cluster. Coordinate European representation in international collaborations spanning multiple continents, so that the European compute and data continuum is presented coherently to partners in WLCG, SRCNet, and adjacent global initiatives, rather than as a federation of separate national positions.

Identity operations (P03). Personal-data baselines (REFEDS DPCoCo and equivalents) are operational platform capabilities rather than project-by-project integrations, with auditing and policy enforcement built in. Federated identity operates as transparent infrastructure across HPC, HTC, and cloud federations.

Community and workforce (P11, P13). Maintain continuous coordination and strategic alignment across the broader network of scientific communities. Permanent computing career paths established across European research institutions through reshaped funding and personnel policies. Long-term plans guarantee sustained support for research software engineering, infrastructure operations, and community coordination roles.

Environmental sustainability (P12). All proposed measures (or adaptations) are consolidated and put into practice. Integrated sustainability policies, accounting, and management operate from project level to infrastructure level, with coordinated supervision by European HEP and RA institutions and EuroHPC/EOSC. Environmental impact budgets are defined by policy makers and enforced. Continuous monitoring of measure effectiveness adapts to the evolution of hardware, software, and the environmental situation.

**End of Phase 3**
Compute and data continuum operates as integrated production capability for HL-LHC and SKA; career paths for research computing and data professionals established; environmental impact reporting operational.

## 6.4. Cross-Priority Dependencies

The priorities are interdependent. Table 6.4.1 shows the primary dependencies that constrain sequencing. A dependency means that progress on the target priority is limited without progress on the source priority.

Table 6.4.1: Key cross-priority dependencies.

| Target Priority | Depends On | Nature of Dependency |
|---|---|---|
| **P02 (Allocation)** | P01 (Governance), P03 (Identity), P09 (Open Science) | Multi-year allocation requires governance agreements, authenticated access, and FAIR data practices |
| **P04 (Compute)** | P01 (Governance), P02 (Allocation), P03 (Identity) | Heterogeneous computing integration requires governance agreements, allocation frameworks, and identity foundations |
| **P05 (Data)** | P01 (Governance), P02 (Allocation), P03 (Identity), P04 (Compute) | Federated data management requires governance agreements, allocation frameworks, authentication, and compute integration |
| **P06 (Workflows)** | P01 (Governance), P02 (Allocation), P03 (Identity), P04 (Compute), P05 (Data), P08 (Portability) | Cross-facility workflows depend on governance, allocation, identity, compute, data, and portable software |
| **P07 (AI/ML)** | P04 (Compute), P05 (Data), P06 (Workflows) | Production AI/ML depends on integrated compute, data, and orchestration |
| **P08 (Portability)** | P04 (Compute), P09 (Open Science), P10 (Preservation), P11 (Community), P13 (Skills) | Portability standards require computing environments, open science practices, preservation, community co-design, and trained workforce |
| **P09 (Reproducibility)** | P04 (Compute), P05 (Data), P08 (Portability), P10 (Preservation), P13 (Skills) | Reproducibility requires compute access, FAIR data, portable software, preserved artefacts, and trained data stewards |
| **P10 (Preservation)** | P04 (Compute), P08 (Portability), P09 (Reproducibility), P13 (Skills) | Preservation requires compute infrastructure, portability standards, FAIR practices, and trained staff |

Priorities P08, P09, and P10 are co-dependent rather than sequential; they must evolve together, with progress in each reinforcing the others.

Pillar 4 priorities (P11, P12, P13) are enabling conditions that span all phases and support progress across all other priorities. P11 (Community Collaboration) provides the co-design mechanisms through which governance decisions (Pillar 1), technical architecture (Pillar 2), and software priorities (Pillar 3) are shaped by scientific community requirements.



## 6.5. Alignment with European Timescales

The roadmap aligns with three external timescales:

Scientific programmes. HL-LHC operations begin in the period 2028-2032 (Run 4 onwards) and ramp through the late 2030s to data volumes requiring the full compute and data continuum. SKA Observatory operations continue beyond 2040. The foundation and integration phases must deliver operational capabilities before these timescales, not after. Delivery in FP10 is therefore critical.

European funding cycles. Horizon Europe (to 2027) supports design and piloting in the foundation phase. FP10 (2028-2034), with research infrastructures in Pillar IV, aligns with the integration phase. The 2028-2034 MFF European Competitiveness Fund may provide a pathway from research to infrastructure deployment. Long-term maturation requires structural commitments that outlast individual framework programmes.

Policy evolution. The forthcoming ERA Act will establish a binding framework for cross-border research cooperation, and the ERA Policy Agenda (2025-2027) creates short-term opportunities for career development and open science. EOSC Federation development, EuroHPC Federation Platform deployment, and ESFRI Roadmap cycles provide institutional milestones that the roadmap should use as coordination points.

The roadmap will adapt to funding outcomes. The dependencies in Section 6.4 are structural; the specific timelines within each phase will be adjusted based on progress.

# 7. Broader Impact



For data-intensive science, a working compute and data continuum changes what large-scale European research does day to day and promises an enormous efficiency gain. A researcher accesses reconstructed data from federated sites across several member states, runs each computational step on the resource best suited to it (e.g. HTC for batch processing, HPC for large-scale simulation, cloud for interactive analysis), and publishes the resulting datasets through shared open archives, all through one identity, one allocation, and one portable software stack. Simultaneously AI will deliver new ways to effectively use and interact with this continuum in a seamless manner.

The investments proposed in this agenda serve this scientific shift first: enabling HL-LHC, SKA, LOFAR, and adjacent programmes to extract their scientific return, and producing methods, software, and infrastructure designs that other data-intensive sciences with similar characteristics identified in Section 1.1. adopt. They also generate impacts on workforce development, the economy, society, and European policy that strengthen the case for sustained public investment.

This section outlines the principal impact pathways, drawing on the RI-PATHS framework (Griniece et al., H2020) for socio-economic impact assessment of research infrastructures. RI-PATHS traces each pathway from Resources and Activity through Outputs and Outcomes to Impacts, and groups thirteen impact pathways under four areas, addressed below in turn: Human Resources (P5, P6), Economy and Innovation (P2, P4, P7), Society (P8, P10, P13), and Policy (P12). Scientific and technological impact is addressed throughout Sections 3 to 6; this section concentrates on the socio-economic pathways that complement them.

## 7.1. Impact on Human Resources

The continuum invests in workforce capacity through P01, P11 and P13 (resources and activity), producing research software engineers, federated-operations engineers, AI/ML specialists, sustainability analysts, and governance and coordination professionals who run cross-infrastructure forums, operate multi-year allocation frameworks, and bridge thematic and horizontal infrastructure governance (outputs). These professionals enable European scientific computing to operate at HL-LHC and SKA scale and sustain the multi-country governance that makes federation work (outcome); the skills they develop transfer to industry, finance, public administration, and adjacent sciences (impact).

Workforce impact is a societal and economic impact in its own right, and is also a precondition for the impacts described in Sections 7.2 and 7.3. Implementing the compute and data continuum requires professionals who combine domain science with computing expertise: research software engineers, specialists in heterogeneous architectures and AI/ML deployment, federated operations engineers, sustainability analysts, and governance and coordination professionals who staff cross-infrastructure forums, operate multi-year allocation frameworks, and bridge thematic and horizontal infrastructure governance. These roles do not yet have well-established career paths in most European research institutions. Priority 1 generates the demand for governance and coordination capacity; P13 develops career frameworks for technical and coordination workforce together; P11 connects these professionals to scientific communities through co-design. The workforce implications extend across the entire agenda. Simultaneously, AI is incorporated in a human-centric manner, compliant with EU regulations and values, to support this workforce and bolster its impact.

The skills developed are transferable. HEP and RA communities have historically trained professionals who move into industry, finance, and technology, carrying expertise in large-scale data processing, statistical analysis, and distributed system management.

The composition of this workforce, by geographic origin, by gender, and by career stage, is a strategic indicator in its own right. Section 7.3 returns to this point with explicit framing, and Section 4.5 (Pillar 4) places reporting commitments under P11 and P13.

Embedding sustainability practices into infrastructure operations (P12) develops a further set of competences, in energy-aware computing, lifecycle assessment, and impact reporting, that are increasingly sought across sectors.

This section addresses workforce capacity and the priorities that build it. Section 7.3 takes up workforce composition (geographic, gender, recognition) within the broader societal-distribution argument.

## 7.2. Impact on Economy and Innovation

Research infrastructures, digital infrastructures, national computing centres, and institutional providers procure hardware, software, and services across these layers of the continuum (resources and activity), generating European technology contracts in processors, interconnects, system integration, storage, and scientific software (outputs). These contracts strengthen industrial capacity in high-performance computing, data services, and scientific software supply (outcome) and feed into European digital competitiveness (impact).

Research infrastructure procurement creates economic multiplier effects. EuroHPC has committed EUR 3.08 billion to supercomputing, stimulating European technology development in processors (RISC-V through the DARE programme at EUR 240 million), interconnects, and system integration. The compute and data continuum extends this value chain to storage, networking, and software services.

Technology transfer flows both ways. Scientific computing has produced widely adopted tools, from ROOT and Geant4 to containerised workflows and federated identity protocols; in adjacent science domains, technology transfer from radio astronomy contributed key patents underpinning Wi-Fi (CSIRO indoor multipath). Open-source software, a defining feature of HEP and RA computing, generates economic value through reduced duplication and shared development costs.

The sovereignty argument (Section 3.2.7) has direct economic implications. Investing in European open technologies retains procurement value within Europe and builds industrial capacity for digital competitiveness. This connection is already operational at the procurement layer: in April 2026 the Commission used its Cloud Sovereignty Framework, which grades providers across eight sovereignty criteria on SEAL (Sovereignty Effectiveness Assurance Levels) from SEAL-0 to SEAL-4, to award up to EUR 180 million in sovereign cloud contracts over six years to four European providers (Post Telecom with OVHCloud and CleverCloud; STACKIT; Scaleway; Proximus with S3NS, Clarence, and Mistral). The European Competitiveness Fund in the 2028-2034 MFF positions the funding instrument that supports this trajectory at programme level, and the forthcoming Cloud and AI Development Act (CAIDA), part of the Commission's Tech Sovereignty package and expected for proposal in Q2 2026, will operationalise it for cloud and AI infrastructure capacity in line with the AI Continent Action Plan cited in Section 3.2.2.

The SRIDA's priorities contribute to this trajectory, particularly P04 (heterogeneous computing), P07 (AI/ML), and P08 (software portability), which build the scientific-computing layer of the European cloud and AI base that CSF assesses and CAIDA is intended to strengthen.

## 7.3. Impact on Society

The continuum operates as a federated multi-country research infrastructure (resources and activity), producing infrastructure, software, and analytical methods for exabyte-scale processing, federated data management, and AI-assisted analysis (outputs). These artefacts apply wherever large-scale data serves public interest, including climate modelling, public health surveillance, environmental monitoring, and disaster response (outcome), reaching these adjacent domains through the same federated channels that already serve the flagship communities (impact).

Open science practices (Priority 9) keep publicly funded outputs (data, software, methods) available for reuse and for public scrutiny. Accountability for public investment in research depends on this availability.

Environmental sustainability (Priority 12) shapes how research computing operates as a public actor. Energy-aware scheduling, lifecycle impact assessment, and sustainability reporting within research infrastructure can set standards that the broader data-centre industry adopts. Operating within defined impact budgets is the practice that makes accountability concrete.

Structurally, the continuum is multi-country by design at both of its layers. Thematic organisations (HEP through HL-LHC and the WLCG; Radio Astronomy through the SKA Observatory and LOFAR) have operated as multinational federations from their inception. Horizontal digital infrastructures (EuroHPC, EOSC, the EGI Federation, and GÉANT) are designed as European endeavours that pool investment, governance, and capacity across member states. The agenda's recommendations on distribution, recognition, and inclusion follow from this dual structure, with P11 and P13 setting out the reporting expectations.

Geographic balance shapes the investment pattern the continuum assumes. On the thematic side, WLCG distributes computing across more than 40 partner countries through national Tier-1 and Tier-2 sites; SKAO and LOFAR-ERIC each operate as IGO- or ERIC-style federations distributing regional centres and station infrastructure across their member states. On the horizontal side, EuroHPC sites are hosted by consortia spanning multiple member states; the EOSC partnership pools in-kind contributions from organisations across the EU; the EGI Federation coordinates compute and data services through national providers; and GÉANT delivers connectivity through the network of National Research and Education Networks. The continuum extends this distributed pattern, including by recognising research software engineering as a career destination accessible from any participating country.

Gender balance in research computing roles remains a challenge across the field. The thematic organisations (CERN, SKAO, LOFAR-ERIC) and the horizontal e-infrastructures (EuroHPC, EOSC, EGI, GÉANT) each publish or aggregate staff-composition data, and EU-level R&I monitoring tracks gender balance across the research workforce, but no consolidated figure currently exists for the share of women in research software engineering and operations roles across the continuum. P11 calls for consolidated annual reporting on geographic and gender distribution of community engagement, drawing on these existing sources; P13 calls for the same on research software engineering positions established under this agenda. Closing the visibility gap is the precondition for closing the participation gap.

## 7.4. Impact on Policy

The continuum's governance, allocation, and trust



mechanisms operate across institutional and national boundaries (resources and activity), producing concrete models for cross-infrastructure governance, multi-year resource allocation, and federated identity (outputs). These models inform EuroHPC, EOSC, and national infrastructure policy (outcome) and provide templates for cross-border resource sharing beyond HEP and RA, including science diplomacy and European positioning in international standards (impact).

This agenda provides evidence-based inputs to European policy. The priorities and roadmap inform investment decisions within EuroHPC, EOSC, and national programmes with the specificity that resource allocation requires.

The governance models in Pillar 1 (multi-year allocation mechanisms, federated identity frameworks, cross-infrastructure coordination) address structural barriers in European research policy. Successfully implemented, they serve as templates for cross-border resource sharing beyond HEP and RA, and align with the European policy environment that shapes research computing: the AI Continent Action Plan (Section 3.2.2), the forthcoming CAIDA expected for proposal in Q2 2026, the Commission's Cloud Sovereignty Framework already operational in EU procurement, and the ERA Act.

Science diplomacy is a further dimension. The HL-LHC and SKA programmes span over 40 countries. European leadership in the computing infrastructure that supports these programmes strengthens Europe's position in international scientific cooperation and its capacity to shape global standards for data management, software sustainability, and federated governance.

# 8. Conclusion

The SRIDA presents thirteen priorities, five investment areas, and a phased multi-annual roadmap for realising the European compute and data continuum. The priorities address gaps identified through systematic analysis of scientific use cases, infrastructure capabilities, and community requirements, aligned with EOSC, EuroHPC, ESFRI, and domain-specific roadmaps.

The Four Pillars framework (Section 4) organises these priorities across governance, architecture, software, and human responsibility. The investment framework (Section 5) maps priorities to funding sources and investment types and names the categories, coordination, sustained software development, and long-term preservation, that traditional infrastructure budgets routinely underestimate. The multi-annual roadmap (Section 6) sequences actions across foundation, integration, and maturation phases. Progress requires advancement across all four pillars together: governance foundations enable technical integration, software development unlocks infrastructure value, workforce capacity sustains all other investments, and managing environmental impact makes them durable.

Realising this requires coordinated action. Scientific communities and research infrastructures as thematic verticals need to  publish multi-year resource plans, seek representation in horizontal digital infrastructure governance, and invest in transitioning AI to production. Digital infrastructures and service providers need to establish liaison mechanisms with thematic research infrastructures, deploy federated identity and standardised interfaces, maintain community engagement, and support environmentally sustainable development and operation. Policy makers and funding bodies need to recognise data-intensive research infrastructures as stakeholders in EuroHPC and EOSC, enable multi-year allocation mechanisms, set and steer environmental sustainability commitments, and fund software preservation, data stewardship, and career frameworks alongside hardware.

The choices made in the next funding cycle will determine whether Europe converts the instruments it has already committed to into the science they were built to produce. The digital continuum at the EU-level is a critical and necessary component to deliver Europe's strategic ambitions and return of investment. The thirteen priorities, the five investment areas, and the three-phase roadmap formulated in this document are the instrument for that conversion. The next step is for all of us  to act upon them.

# 9. References



1. AARC BPA 2025. AARC-G080: AARC Blueprint Architecture 2025. Available: https://aarc-community.org/guidelines/aarc-g080/
2. AARC Policy Development Kit. AARC Policy Development Kit (AARC-PDK). Available: https://aarc-community.org/policy/policy-development-kit
3. AARC Tree Project. Available: https://aarc-community.org/aarc-tree-project/
4. Aartsen et al. 2014 (IceProd). Aartsen, M. G. et al. (2014). The IceProd Framework: Distributed Data Processing for the IceCube Neutrino Observatory. Available: https://confluence.egi.eu/display/SPECTRUMCoP/The+IceProd+Frame-work%3A+Distributed+Data+Processing+for+the+IceCube+Neutrino+Observatory
5. Abdurachmanov et al. 2014 (ARM/Xeon). Abdurachmanov, D. et al. (2014). Explorations of the viability of ARM and Xeon Phi for physics processing. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Explorations+of+the+-viability+of+ARM+and+Xeon+Phi+for+physics+processing
6. Acosta et al. 2023 (CMS-BSC). Acosta, C. et al. (2023). Integration of the Barcelona Supercomputing Center for CMS computing: Towards large-scale production [Presentation]. Available: https://confluence.egi.eu/display/SPECTRUM-CoP/Integration+of+the+Barcelona+Supercomputing+Center+for+CMS+computing%3A+Towards+large+scale+pro-duction
7. Adelman-McCarthy et al. 2023. Adelman-McCarthy, J. et al. (2023). Extending the distributed computing infrastruc-ture of the CMS experiment with HPC resources [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Extending+the+distributed+computing+infrastructure+of+the+CMS+experiment+with+HPC+resour-ces
8. ADORE Software Declaration. Research Software Alliance, Amsterdam Declaration on Funding Research Software Sustainability (1.1), 2024. Zenodo. Available: https://doi.org/10.5281/zenodo.13735888
9. AI Continent Action Plan. European Commission, "The AI Continent Action Plan", COM(2025) 165. Available: https://digital-strategy.ec.europa.eu/en/library/ai-continent-action-plan
10. AI4ES (US DOE). U.S. Department of Energy, "AI for Energy and Science Report", 2023.
11. ALICE O2 evolution 2019. ALICE Collaboration (2019). Evolution of the O2 system [Technical Report]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/The+ALICE+Collaboration%3A+Evolution+of+the+O2+system
12. Alice Recoque. Available: https://www.eurohpc-ju.europa.eu/contract-signed-alice-recoque-euro-pes-new-exascale-supercomputer-2025-11-18_en
13. Andreozzi e-IRG 2025. S. Andreozzi, "AI Needs in Compute/Data Intensive Research from HEP and RA — The SPECTRUM Project Perspective", e-IRG Workshop, June 2025.
14. Arrhenius. Available: https://www.naiss.se/news/2024/11/request-for-proposals-for-new-supercomputer-arrhe-nius
15. Atif et al. 2023. Atif, M. et al. (2023). Evaluating Portable Parallelization Strategies for Heterogeneous Architectu-res in High Energy Physics. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Evaluating+Portable+Paralleli-zation+Strategies+for+Heterogeneous+Architectures+in+High+Energy+Physics
16. ATLAS 2024 (Google Cloud TCO). ATLAS Collaboration (2024). Total cost of ownership and evaluation of Google cloud resources for the ATLAS experiment at the LHC. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Total+cost+of+ownership+and+evaluation+of+Google+cloud+resources+for+the+ATLAS+experiment+at+the+LHC
17. ATLAS HL-LHC Roadmap. ATLAS Collaboration (2022). ATLAS Software and Computing HL-LHC Roadmap. Available: https://confluence.egi.eu/display/SPECTRUMCoP/ATLAS+Software+and+Computing+HL-LHC+Roadmap
18. Apptainer. https://apptainer.org/
19. Bagnasco et al. 2023 (ET Computing). Bagnasco, S. et al. (2023). Computing Challenges for the Einstein Tele-scope Project [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Computing+Challen-ges+for+the+Einstein+Telescope+Project
20. Bagnasco 2024 (LVK Computing). Bagnasco, S. for the Virgo Collaboration and the LIGO Scientific Collabora-tion (2024). The LIGO-Virgo-KAGRA Computing Infrastructure for Gravitational-wave Research. Available: https://confluence.egi.eu/display/SPECTRUMCoP/The+Ligo-Virgo-KAGRA+Computing+Infrastructure+for+Gravitational-wa-ve+Research
21. Banerjee et al. 2023 (HECAP+). Banerjee, S. et al. (2023). Environmental sustainability in basic research: a perspective from HECAP+ [Technical Report]. Available: https://confluence.egi.eu/pages/viewpage.action?pa-geId=288784649
22. BDVA DataLabs. BDVA DataLabs Initiative. Available: https://bdva.eu/task-forces/bdva-data-labs-initiative/
23. Beche et al. 2018. Beche, A. et al. (2018). Supercomputers, Clouds and Grids powered by BigPanDA for Brain studies [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Supercomputers%2C+Cl-ouds+and+Grids+powered+by+BigPanDA+for+Brain+studies
24. Bertone et al. 2024 (NP Tools). Bertone, V. et al. (2024). Nuclear Physics Tools — Machine Learning, Artificial In-telligence, and Quantum Computing [Report]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Nuclear+P-hysics+Tools+%E2%80%93+Machine+Learning%2C+Artificial+Intelligence%2C+and+Quantum+Computing
25. Beyer et al. 2020 (DESY Photon). Beyer, C. et al. (2020). Beyond HEP: Photon and accelerator science compu-ting infrastructure at DESY [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Beyo-

nd+HEP%3A+Photon+and+accelerator+science+computing+infrastructure+at+DESY
**26.** Boccali 2019 (Computing Models HEP). Boccali, T. (2019). Computing models in high energy physics. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Computing+models+in+high+energy+physics
**27.** Boccali et al. 2019 (Cross-experiment HPC). Boccali, T. et al. (2019). Summary of the cross-experiment HPC workshop [Workshop Summary]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Summary+of+the+-cross-experiment+HPC+workshop
**28.** Boccali et al. 2021 (INFN-CINECA). Boccali, T. et al. (2021). Enabling HPC Systems for HEP: The INFN-CINECA Experience. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Enabling+HPC+Systems+for+HEP%3A+The+-INFN-CINECA+Experience
**29.** Böhler et al. 2021 (Opportunistic WLCG). Böhler, M. et al. (2021). Transparent Integration of Opportunistic Resources into the WLCG Compute Infrastructure [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Transparent+Integration+of+Opportunistic+Resources+into+the+WLCG+Compute+Infrastructure
**30.** Boveia et al. 2024. Boveia, A. et al. (2024). Snowmass 2021 Cross Frontier Report: Dark Matter Complementarity (Extended Version). Available: https://confluence.egi.eu/pages/viewpage.action?pageId=288785184
**31.** Breen et al. 2021 (SKAO SDP). Breen, S. et al. (2021). SKAO Science Data Products: A Summary [Data Product Summary]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/SKAO+Science+Data+Products%3A+A+Summary
**32.** Britton 2024 (Carbon Cost). Britton, D. (2024). Simulating the Carbon Cost of Grid Sites [Presentation]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Simulating+the+Carbon+Cost+of+Grid+Sites
**33.** Britton, Campana, Panzer-Stradel 2024 (WLCG Energy). Britton, D., Campana, S. & Panzer-Stradel, B. (2024). A holistic study of the WLCG energy needs for the LHC scientific program [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/A+holistic+study+of+the+WLCG+energy+needs+for+the+LHC+scientific+program
**34.** Bruers et al. 2023. Bruers, B. et al. (2023). Resource-aware Research on Universe and Matter: Call-to-Action in Digital Transformation [Meeting Summary]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Resource-aware+Research+on+Universe+and+Matter%3A+Call-to-Action+in+Digital+Transformation
**35.** Casper et al. 2022 (Small HEP). Casper, D. et al. (2022). Software and Computing for Small HEP Experiments [Workshop Summary]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Software+and+Computing+for+Small+HEP+Experiments
**36.** CernVM-FS. Available: https://cernvm.cern.ch/fs/
**37.** Chue Hong et al. 2011. Chue Hong, N. P., Aragon, S., Hettrick, S., & Jay, C. (2011). The future of research software is the future of research. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC12416084/
**38.** Cloud Container Platform. Available: https://open-science-cloud.ec.europa.eu/services/cloud-container-platform
**39.** CMS HPC Integration 2020. CMS Offline Software and Computing (2020). HPC Resources Integration at CMS. Available: https://confluence.egi.eu/display/SPECTRUMCoP/HPC+Resources+Integration+at+CMS
**40.** CMS Phase-2 Computing Model. CMS Offline Software and Computing (2022). CMS Phase-2 Computing Model: Update. Available: https://confluence.egi.eu/display/SPECTRUMCoP/CMS+Phase-2+Computing+Model%3A+Update?src=contextnavpagetreemode
**41.** Copernicus Earth-observation data space. Available: https://dataspace.copernicus.eu/
**42.** Couvares 2021 (3G GW). Couvares, P. (2021). Gravitational-Wave Data Analysis: Computing Challenges in the 3G Era [Technical Report]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Gravitational-Wave+Data+Analysis%3A+Computing+Challenges+in+the+3G+Era
**43.** CSCS. Available: https://www.cscs.ch/services/overview
**44.** D3.1. SPECTRUM Project, "D3.1 Community of Practice — First Period Report", 2024. Available: https://zenodo.org/records/15100586
**45.** D4.1. Community of Practice and sustainability plan - Final report. To be published in June 2026
**46.** D5.1. SPECTRUM Project, "D5.1 Representative Use Cases: Analysis and Alignment", 2024. Available: https://zenodo.org/records/15310223
**47.** D5.2. SPECTRUM Project, "D5.2 Interoperable Access Policies: Analysis and Recommendations", 2024. Available: https://zenodo.org/records/15647608
**48.** D5.3. SPECTRUM Project, "D5.3 Landscape of Research Infrastructures: Technologies, Services, Gaps", 2024. Available: https://zenodo.org/records/15647755
**49.** D6.1. SPECTRUM Project, "D6.1 Technical Blueprint for Compute and Data Continuum", 2026. https://zenodo.org/records/20138626
**50.** Daedalus. Available: https://www.eurohpc-ju.europa.eu/daedalus-call-acquire-new-european-world-class-supercomputer-greece-2024-06-25_en
**51.** Deucalion. Available": https://rnca.fccn.pt/en/deucalion/
**52.** Dewdney et al. 2022 (SKA1). Dewdney, P. et al. (2022). SKA1 Design Baseline Description [Design Baseline Description]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/SKA1+Design+Baseline+Description



**53.** Diefenthaler et al. 2022. Diefenthaler, M. et al. (2022). Future Trends in Nuclear Physics Computing [Editorial]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Future+Trends+in+Nuclear+Physics+Computing
**54.** Digital Decade 2030. European Commission, "Digital Decade Policy Programme 2030", 2021. Decision (EU) 2022/2481 of the European Parliament and of the Council of 14 December 2022 establishing the Digital Decade Policy Programme 2030. ELI: http://data.europa.eu/eli/dec/2022/2481/oj. Available: https://digital-strategy.ec.europa.eu/en/library/digital-decade-policy-programme-2030
**55.** Discoverer. Available:https://sofiatech.bg/en/petascale-supercomputer/
**56.** DORA. Allen, L., Barbour, V., Cobey, K., Faulkes, Z., Hazlett, H., Lawrence, R., Lima, G., Massah, F., & Schmidt, R. (2025). A Practical Guide to Implementing Responsible Research Assessment at Research Performing Organizations. Declaration on Research Assessment (DORA). Available: https://sfdora.org/
**57.** DUNE Computing CDR. The DUNE Collaboration (2022). DUNE Offline Computing Conceptual Design Report [Technical Report]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/DUNE+Offline+Computing+Conceptual+Design+Report
**58.** EBRAINS neuroscience platform. Available: https://www.ebrains.eu/
**59.** EGI 2024 Annual Report. EGI Federation, "EGI Federation Annual Report 2024", 2024. Available: https://www.egi.eu/
**60.** EGI EOSC Federation Discussion Paper. EGI Federation, "EGI Contribution to the EOSC Federation — Discussion Paper v1.2", May 2024. Available: https://www.egi.eu/
**61.** EGI FedCloud. Available: https://www.egi.eu/egi-infrastructure/#layers-federated-cloud
**62.** EGI HTC-oriented infrastructure. Available: https://www.egi.eu/egi-infrastructure/#layers-federated-htc
**63.** eIDAS 2.0. European Commission, "eIDAS 2.0. European Identity Framework". Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC. ELI: http://data.europa.eu/eli/reg/2014/910/oj. Available: https://digital-strategy.ec.europa.eu/en/policies/eidas-regulation
**64.** Elvira et al. 2022. Elvira, V. D. et al. (2022). The Future of High Energy Physics Software and Computing: Report of the 2021 US Community Study on the Future of Particle Physics. Available: https://confluence.egi.eu/display/SPECTRUMCoP/The+Future+of+High+Energy+Physics+Software+and+Computing%3A+Report+of+the+2021+US+Community+Study+on+the+Future+of+Particle+Physics
**65.** EOSC Federation. Available: https://eosc.eu/building-the-eosc-federation/
**66.** EOSC SRIA. European Open Science Cloud, "EOSC Strategic Research and Innovation Agenda v1.3", November 2024. Available: https://eosc.eu/sria
**67.** EPCC (UK). Available: https://www.epcc.ed.ac.uk/
**68.** EP Resolution on ERA Act. European Parliament, "Resolution of 10 March 2026 on the upcoming European Research Area (ERA) Act", P10_TA(2026)0068, 2026. Available: https://www.europarl.europa.eu/doceo/document/TA-10-2026-0068_EN.html
**69.** ErUM4 Data Hub. Available: https://erumdatahub.de/en/
**70.** ESCAPE Project's Open-source Scientific software and Service Repository. Available: https://projectescape.eu/services/open-source-scientific-software-and-service-repository-ossr
**71.** ESFRI Landscape Analysis 2024. European Strategy Forum on Research Infrastructures, "ESFRI Landscape Analysis 2024", 2024. Available: https://www.esfri.eu/
**72.** ESFRI Position Paper on ERA Act. ESFRI, "Position Paper on the European Research Area Act (ERA Act)", FI26-WP-20260115, January 2026.
**73.** ESFRI Roadmap 2021. European Strategy Forum on Research Infrastructures, "ESFRI Roadmap 2021", 2021. Available: https://www.esfri.eu/
**74.** ETP4HPC SRA6. European Technology Platform for High Performance Computing, "ETP4HPC Strategic Research Agenda 6", 2024. Available: https://www.etp4hpc.eu/
**75.** European Interoperability Framework. European Commission, "European Interoperability Framework — Implementation Strategy", COM(2017) 134 final, 2017. Updated under Regulation (EU) 2024/903 (Interoperable Europe Act), 2024. Available: https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/european-interoperability-framework-detail
**76.** European Particle Physics Strategy 2020. European Particle Physics Strategy, "2020 Update of the European Strategy for Particle Physics", 2020.
**77.** EU Data Strategy 2020. European Commission, "European Data Strategy", 2020. Available: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/european-data-strategy
**78.** EU Data Union Strategy. European Commission, "European Data Union Strategy". Available: https://digital-strategy.ec.europa.eu/en/policies/data-union
**79.** EuroHPC Federation Platform. EuroHPC Federation Platform. Available: https://www.eurohpc-ju.europa.eu/supercomputers/eurohpc-federation-platform_en
**80.** EuroHPC Joint Undertaking (JU). Available:https://www.eurohpc-ju.europa.eu/index_en
**81.** EuroHPC MAWP 2024. EuroHPC Joint Undertaking, "EuroHPC Multi-Annual Work Programme", 2024.

82. EuroHPC Strategic Agenda. EuroHPC Joint Undertaking, “EuroHPC Strategic Agenda 2024”, 2024. Available: https://eurohpc-ju.europa.eu/
83. EuroHPC Summit 2024. EuroHPC (2024). EuroHPC Summit 2024: Interconnecting EuroHPC Supercomputers for Scientific and Industrial Advancement [Presentation]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/EuroHPC+Summit+2024%3A+Interconnecting+EuroHPC+Supercomputers+for+Scientific+and+Industrial+Advancement
84. EVERSE Software. Available: https://everse.software
85. EuroHPC Supercomputers. Available: https://www.eurohpc-ju.europa.eu/supercomputers/our-supercomputers_en
86. EVERSE. EVERSE Project. Psomopoulos, F., Capella-Gutierrez, S., Portell-Silva, L., & Pechlivanis, N. (2024, January 18). EOSC-EVERSE: Paving the way towards a European Virtual Institute for Research Software Excellence. Zenodo. Available: https://doi.org/10.5281/zenodo.10526785
87. EXTRACT. EXTRACT Project. Available: https://extract-project.eu/
88. EXTRACT-TASKA (Cecconi). Cecconi, B., Le Sidaner, P. & Aicardi, S. (2026). EXTRACT-TASKA: A Distributed Workflow Orchestrator for Radio Astronomy Data Processing. EOSC Winter School 2026, Nice. Zenodo. Available: https://doi.org/10.5281/zenodo.19093528
89. FAIR4RS. FAIR for Research Software (FAIR4RS): A summary, August 2022. Available: https://www.researchsoft.org/blog/2022-08/
90. Federated Software Catalogue. Available: https://www.eessi.io/docs/blog/2026/03/09/EESSI-in-EFP-update/#eessi-as-a-base-for-the-efp-federated-software-catalog
91. Filipcic, Girone & Boccali 2024 (HEP-HPC Europe). Filipcic, A., Girone, M. & Boccali, T. (2024). HEP/HPC Strategy Meeting — Europe [Meeting Summary]. Available: https://confluence.egi.eu/pages/viewpage.action?pageId=298322551
92. Fischer et al. 2020 (Heterogeneous). Fischer, M. et al. (2020). Effective Dynamic Integration and Utilization of Heterogeneous Compute Resources [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Effective+Dynamic+Integration+and+Utilization+of+Heterogenous+Compute+Resources
93. Fischer et al. 2020 (Opportunistic). Fischer, M. et al. (2020). Lightweight dynamic integration of opportunistic resources [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Lightweight+dynamic+integration+of+opportunistic+resources
94. Fisk et al. 2024 (HEP-HPC USA). Fisk, I. et al. (2024). HEP/HPC Strategy Meeting — USA [Meeting Summary]. Available: https://confluence.egi.eu/pages/viewpage.action?pageId=298322541
95. Gal et al. 2019 (KM3NeT). Gal, T. et al. (2019). KM3NeT Report on Documentation Strategy, Environment, and Software [Deliverable]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/KM3NeT+Report+on+Documentation+Strategy%2C+Environment%2C+and+Software
96. GANIL Data Policy 2020. GANIL (2020). GANIL Data Policy [Technical Report]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/GANIL+Data+Policy
97. GCS. Available: https://www.gauss-centre.eu/for-users/hpc-access
98. GENCI. Available: http://www.idris.fr/eng/info/gestion/demandes-heures-eng.html
99. Girone & Boccali 2025 (HEP-HPC All Regions). Girone, M. & Boccali, T. (2025). HEP/HPC Strategy Meeting — All Regions [Meeting Summary]. Available: https://confluence.egi.eu/pages/viewpage.action?pageId=298322557
100. Girone 2019 (HPC for LHC). Girone, M. (2019). Common challenges for HPC integration into LHC computing [Technical Report]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Common+challenges+for+HPC+integration+into+LHC+computing
101. GSI/FAIR OS Licence Guidelines. GSI/FAIR Collaboration (2021). Open source software licences at GSI/FAIR — Guidelines. Available: https://confluence.egi.eu/pages/viewpage.action?pageId=288785215
102. GSI/FAIR Upload Instructions. GSI/FAIR Collaboration (2023). Instructions for uploading and linking research data and software at GSI Helmholtzzentrum für Schwerionenforschung GmbH. Available: https://confluence.egi.eu/pages/viewpage.action?pageId=288785219
103. HEPSF HL-LHC Review 2020. HEP Software Foundation (2020). HL-LHC Computing Review: Common Tools and Community Software [Technical Report]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/HL-LHC+Computing+Review%3A+Common+Tools+and+Community+Software
104. HEP-SF Strategic Plan 2026. High Energy Physics Software Foundation, “HEP Software Foundation Strategic Plan 2026”, 2023.
105. HEP-SF Roadmap. High Energy Physics Software Foundation, “HEP Software Foundation Roadmap”, 2022.
106. HLRS. Available: https://www.hlrs.de/
107. Hollowell et al. 2023. Hollowell, C. et al. (2023). Financial Case Study: Use of Cloud Resources in HEP Computing [Presentation, CHEP 2023]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Financial+Case+Study%3A+Use+of+Cloud+Resources+in+HEP+Computing
108. HyperCON Final Report. EuroHPC Joint Undertaking, “HyperCON Final Report”, 2023.
109. ICSC. Available: https://www.supercomputing-icsc.it/
110. IHDEA (international heliophysics data environment alliance. Available: https://ihdea.net
111. InterTwin. interTwin Project. Available: https://www.intertwin.eu/
112. IVOA (international virtual observatory alliance. Available: https://ivoa.net
113. IPDA, International Planetary Data Alliance. Available: https://planetarydata.org
114. JSC. Available: https://www.fz-juelich.de/en/jsc
115. JENA Computing White Paper. Joint ECFA-NuPECC-APPEC Activities (JENA), “Computing White Paper”, 2023.
116. JENA WG1 (HPC). JENA Working Group 1, “HPC for Fundamental Science”, 2023.
117. JENA WG2 (Software). JENA Working Group 2, “Software and Applications”, 2023.
118. JENA WG3 (Data). JENA Working Group 3, “Data Management and Access”, 2023.
119. JENA WG4 (Infrastructure). JENA Working Group 4, “Infrastructure and Facilities”, 2023.
120. JENA WG5 (Federation). JENA Working Group 5, “Federation and Interoperability”, 2023.
121. JUPITER. Available: https://www.fz-juelich.de/en/ias/jsc/jupiter
122. Kahn et al. 2022 (Dark Matter Snowmass). Kahn, Y. et al. (2022). Snowmass2021 Cosmic Frontier: Modeling, statistics, simulations, and computing needs for direct dark matter detection. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Snowmass2021+Cosmic+Frontier%3A+Modeling%2C+statistics%2C+simulations%2C+and+computing+needs+for+direct+dark+matter+detection
123. Karolina. Available: https://www.it4i.cz/en/infrastructure/karolina
124. Kokkos. Available: https://github.com/kokkos/kokkos
125. Laycock et al. 2024 (Einstein Telescope). Laycock, P. et al. (2024). Preparatory Phase for the Einstein Telescope Gravitational Wave Observatory: Computing and Data Requirements [Deliverable]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Preparatory+Phase+for+the+Einstein+Telescope+Gravitational+Wave+Observatory%3A+Computing+and+Data+Requirements
126. Leonardo. Available: https://leonardo-supercomputer.cineca.eu/
127. LOFAR ERIC Access Policy. Interim LOFAR ERIC Council (2021). LOFAR ERIC Access Policy to Scientific User Services [Policy Document]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/LOFAR+ERIC+Access+Policy+to+Scientific+User+Services
128. LOFAR ERIC Science Data Policy. LOFAR ERIC Council (2023). Science Data Policy of LOFAR ERIC [Policy Document]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Science+Data+Policy+of+LOFAR+ERIC
129. LOFAR Long-Term Archive (LTA). Available: https://lta.lofar.eu/
130. LOFAR2.0 Summary 2023. LOFAR (2023). LOFAR2.0 compared to LOFAR: a short summary [Technical Summary]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/LOFAR2.0+compared+to+LOFAR%3A+a+short+summary
131. LRZ. Available: https://www.lrz.de/en/
132. LUMI. Available: https://www.lumi-supercomputer.eu/about-lumi/
133. Maeno et al. 2024 (PanDA in ATLAS). Maeno, T. et al. (2024). Utilizing Distributed Heterogeneous Computing with PanDA in ATLAS [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Utilizing+Distributed+Heterogeneous+Computing+with+PanDA+in+ATLAS
134. MareNostrum 5. Available: https://www.bsc.es/marenostrum/marenostrum-5
135. Megino et al. 2022 (US ATLAS/CMS HPC). Barreiro Megino, F. et al. (2022). US ATLAS and US CMS HPC and Cloud Blueprint [Technical Report]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/US+ATLAS+and+US+CMS+HPC+and+Cloud+Blueprint
136. Meluxina. Available: https://www.luxprovide.lu/meluxina/
137. MultiXscale . Available: https://www.multixscale.eu/
138. NHR Alliance. Available: https://www.nhr-verein.de/en/computing-time
139. NuPECC TWG9. NuPECC TWG9 (2024). Open Science and Data. Available: https://confluence.egi.eu/display/SPECTRUMCoP/NuPECC+TWG9+%3A+Open+Science+and+Data
140. ODISSEE project. Available: https://odissee-project.eu/
141. OSCARS. OSCARS Project. Nicoletta, C., Carrillo, R., Fiser, D., Wolff-Boenisch, B., Schmidt-Tremmel, F., Lamanna, G., Petzold, A., & Hienola, A. (2023). OSCARS project funded to foster the uptake of Open Science in Europe. Zenodo. Available: https://doi.org/10.5281/zenodo.10379119
142. Physics Briefing Book 2020. European Strategy for Particle Physics (2020). Physics Briefing Book: Input for the European Strategy for Particle Physics Update 2020 [Strategy Report]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Physics+Briefing+Book+%3A+Input+for+the+European+Strategy+for+Particle+Physics+Update+2020
143. Physics Briefing Book 2026. Physics Briefing Book: Input for the 2026 update of the European Strategy for Particle Physics, 2025. Available: https://arxiv.org/abs/2511.03883, https://cds.cern.ch/record/2944678/files/Physics%20Briefing%20Book.pdf
144. Pérez-Calero Yzquierdo 2020. Pérez-Calero Yzquierdo, A. (2020). CMS strategy for HPC resource exploitation [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/CMS+strategy+for+HPC+resource+exploitation
145. Pizzo et al. 2023 (LOFAR2.0 DM). Pizzo, R. et al. (2023). LOFAR2.0 Data Management Capabilities [Note]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/LOFAR2.0+Data+Management+Capabilities
146. PUNCH4NFDI. Available: https://www.punch4nfdi.de/

147. Reppin et al. 2021 (DESY Jupyter). Reppin, J. et al. (2021). Interactive Analysis Notebooks on DESY Batch Resources: Bringing Jupyter to HTCondor and Maxwell at DESY. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Interactive+Analysis+Notebooks+on+DESY+Batch+Resources%3A+Bringing+Jupyter+to+HTCondor+and+Maxwell+at+DESY
148. RES. Available: https://www.res.es/en/access-to-res
149. Roberts et al. 2019 (DANCE). Roberts, A. et al. (2019). Dark-matter And Neutrino Computation Explored (DANCE) Community Input to Snowmass [Meeting Summary]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Dark-matter+And+Neutrino+Computation+Explored+%28DANCE%29+Community+Input+to+Snowmass
150. Rohr & Eulisse 2023 (ALICE O2 GPU). Rohr, D. & Eulisse, G. (2023). The O2 software framework and GPU usage in ALICE online and offline reconstruction in Run 3 [Presentation]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/The+O2+software+framework+and+GPU+usage+in+ALICE+online+and+offline+reconstruction+in+Run+3
151. Salgado et al. 2023 (SRCNet). Salgado, J. et al. (2023). SKA Regional Centres Network (SRCNet) Software Architecture Document [Design Description Document]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/SKA+-Regional+Centres+Network+%28SRCNet%29+Software+Architecture+Document
152. SKA SRCNet. Available: https://www.skao.int/en/science-users/119/ska-regional-centres
153. SCIENCE Clusters. SCIENCE Clusters. Available: https://science-clusters.eu/
154. Sfiligoi et al. 2014. Sfiligoi, I. et al. (2014). CMS experience of running glideinWMS in High Availability mode [Conference Paper]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/CMS+experience+of+running+glideinWMS+in+High+Availability+mode
155. Simili et al. 2024 (ARM vs x86). Simili, E. et al. (2024). ARMing HEP for the future: Energy Efficiency of WLCG sites (ARM vs. x86) [Conference Paper]. Available: https://confluence.egi.eu/pages/viewpage.action?pageId=302876052
156. Simpl Data Federation Platform. Available: https://digital-strategy.ec.europa.eu/en/policies/simpl
157. SMP Template (Martinez-Ortiz). Martinez-Ortiz, C., Martinez Lavanchy, P., Sesink, L., Olivier, B. G., Meakin, J., de Jong, M., & Cruz, M. (2023). Practical guide to Software Management Plans (1.1). Zenodo. Available: https://doi.org/10.5281/zenodo.7589725
158. Spider. Available: https://servicedesk.surf.nl/wiki/spaces/WIKI/pages/9797664/Data+Processing+-+Spider
159. SRCSC WG6 2021. SKA Regional Centre Steering Committee (SRCSC) WG6, Task Package 1 (2021). SKA Science Regional Centres Community Input Questionnaire [Survey]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/SKA+Science+Regional+Centres+Community+Input+Questionnaire
160. Stagni 2023 (DiracGrid). Stagni, F. (2023). Status of DiracGrid projects [Presentation]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/Status+of+DiracGrid+projects
161. SURF's Grid. Available: https://servicedesk.surf.nl/wiki/spaces/WIKI/pages/1474704/Data+Processing+-+GRID
162. Vega. Available: https://izum.si/en/vega-en/
163. Wagg et al. 2021. Wagg, J. et al. (2021). SKA1 Scientific Use Cases [Technical Document]. Available: https://confluence.egi.eu/display/SPECTRUMCoP/SKA1+Scientific+Use+Cases
164. WLCG Strategy. Worldwide LHC Computing Grid (WLCG), "WLCG Strategy Document 2027", 2023. Available: https://wlcg.web.cern.ch/



# Annexes

You can consult the full annexes in the online version of the SRIDA: 10.5281/zenodo.20182879

Annex A: Priorities

- Pillar 1: Policy, Trust & Governance
- Pillar 2: Architecture & Interoperability
- Pillar 3: Software & Science Enablers
- Pillar 4: Human Capital & Responsibility

Annex B: Priority Description Template

Annex C: Priority Evidence and Validation

- C.1 Desk Research and Evidence Sources
- C.2 Feedback & Alignment from External events
- C.3 Community of Practice and Advisory Board Validation
- C.4 Priority to Technical Blueprint Mapping

## Terminology / Acronyms

| Terminology / Acronym | Definition |
|---|---|
| AI | Artificial Intelligence |
| AARC BPA | AARC Blueprint Architecture |
| CAIDA | Cloud and AI Development Act |
| CoP | Community of Practice |
| CPU | Central Processing Unit |
| CUE | Carbon Usage Effectiveness |
| DB | Database |
| DOI | Digital Object Identifier |
| DPU | Data Processing Unit |
| EAB | External Advisory Board |
| EOSC | European Open Science Cloud |
| ERA | European Research Area |
| ERDF | European Regional Development Fund |
| ESA | European Space Agency |
| ESCAPE | European Science Cluster of Astronomy and Particle Physics ESFRI research infrastructures |
| ESFRI | European Strategy Forum on Research Infrastructures |
| FAIR | Findable, Accessible, Interoperable, Reusable |
| FIAM | Federated Identity Access Management |
| FPGA | Field-Programmable Gate Array |
| GDPR | General Data Protection Regulation |
| GPU | Graphics Processing Unit |
| HEP | High Energy Physics |
| JENA | Joint ECFA-NuPECC-APPEC |
| LHC | Large Hadron Collider |
| HL-LHC | High Luminosity Large Hadron Collider |
| HTC | High-Throughput Computing |



| | |
|---|---|
| HPC | High Performance Computing |
| KER | Key Exploitable Result |
| KPI | Key Performance Indicator |
| LOFAR | LOw Frequency ARray |
| MFA | Multi-factor authentication |
| ML | Machine Learning |
| MoU | Memorandum of understanding |
| NGI | National federation of shared computing, storage and data resources part of EGI |
| NREN | National Research and Education Network |
| PUE | Power Usage Effectiveness |
| RA | Radio Astronomy |
| RI | Research Infrastructure |
| SKA | Radio Telescopes managed by SKAO |
| SKAO | SKA Observatory: Intergovernmental organisation |
| SLA | Service Level Agreement |
| SRCNet | SKA Regional Centre Network |
| SRIDA | Strategic Research, Innovation and Deployment Agenda |
| SSO | Single Sign On |
| TI | Technology Infrastructure |
| VA/TNA | Virtual Access / Transnational Access |
| WLCG | Worldwide LHC Computing Grid |
| SSO | Single Sign On |
| VO | Virtual Observatory |
| VOMS | Virtual Organisation Membership Service |
| VM | Virtual Machine |
| WG | Working Group |
| WLCG | Worldwide LHC Computing Grid |
| WP | Work Package |

SPECTRUM

Strategic Research Innovation and Deployment Agenda (SRIDA)

www.spectrumproject.eu

Funded by the European Union

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