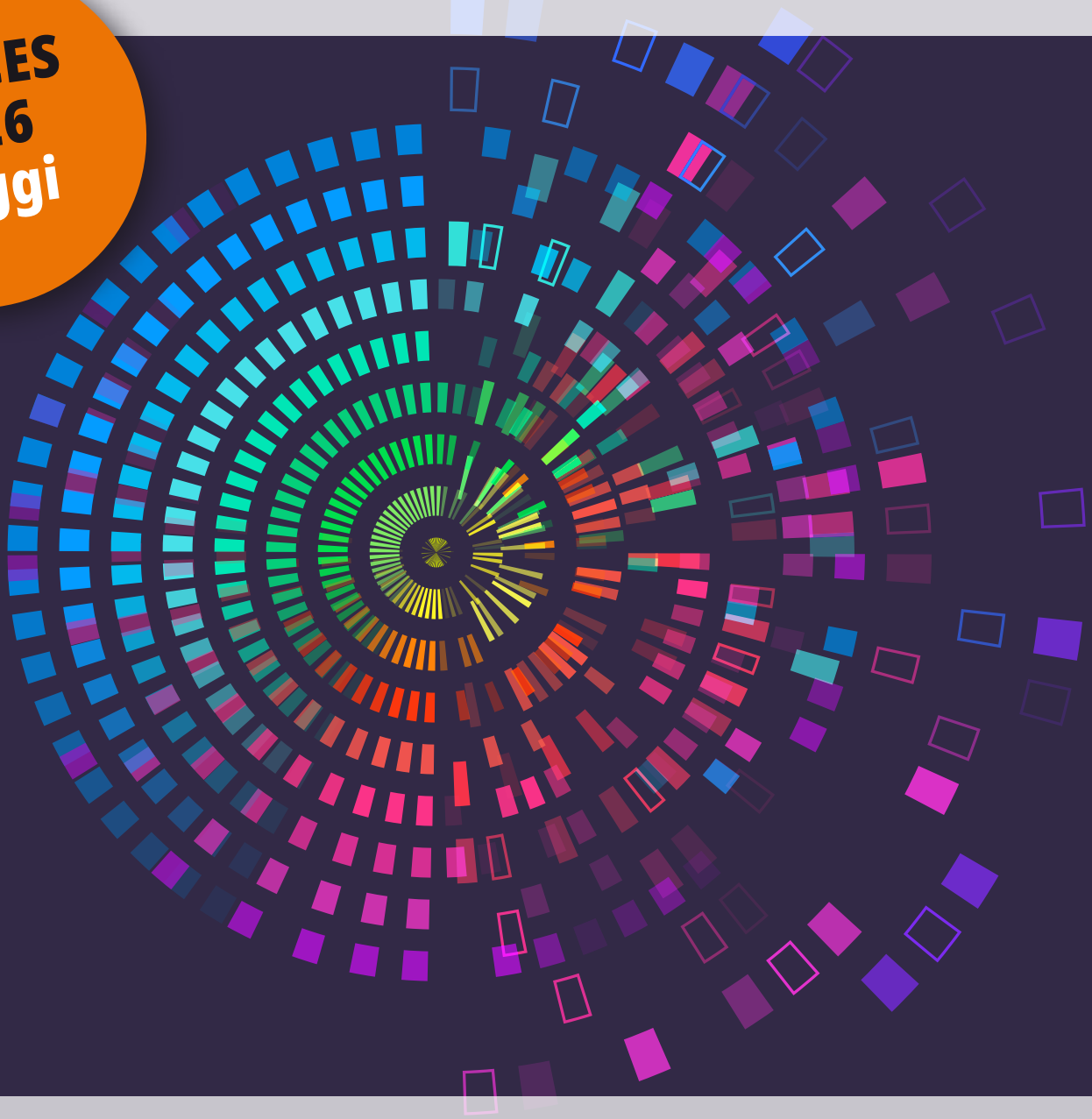
HiPEAC info 78

JUNE 2026

ACACES 2026 Fiuggi



The policies, architectures, and compute spinning raw data into value

HiPEAC startups gaining momentum

When AI meets safety engineering: Locating enforcement in the stack

Contents

How BDVA helps extract value from data

HiPEAC Vision 2026

The HiPEAC companies gaining momentum

Tailoring the compute to the data: Eduardo Quiñones presents EXTRACT

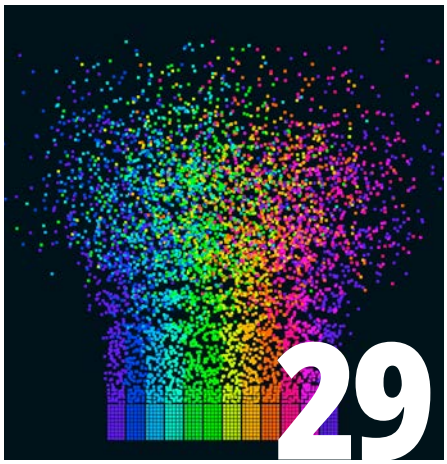


Photo credit: Oleh | stock.adobe.com

Wrestling data from edge to cloud, with the HiPEAC community

Opinion: What DoD vs Anthropic means for safety engineering

Spanning the compute continuum from edge to cloud, HiPEAC (High Performance, Edge And Cloud computing) is a network of over 2,500 world-class computing systems researchers, industry representatives and students. First established in 2004, the project is now in its eighth edition. HiPEAC8 focuses on networking, roadmapping and innovation activities: bringing the computing community together in Europe, exchanging ideas, inspiring entrepreneurs and exploring the long-term vision for computing systems.

hipeac.net

hipeac.net/linkedin

@hipeac / @hipeacjobs

hipeac.net/tv

Funded by the European Union

The HiPEAC project has received funding from the European Union's Horizon Europe research and innovation funding programme under grant agreement number 101069836. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Cover image: arthead | stock.adobe.com

Design: www.magelaan.be

Editor: Madeleine Gray

This magazine marks the end of HiPEAC7 and the start of HiPEAC8. Not everybody is aware that HiPEAC is not only a community, but also a Horizon Europe project. Without the funding provided by the project, HiPEAC would not be able to serve the community as it has been doing since 2004, with services now including yearly vision documents, a conference, a summer school, a jobs website, and communication and dissemination support. I am happy to announce that we have been successful in securing funding for the next three years, until 2029, which also marks the 25th anniversary of the HiPEAC community. We are very grateful for the years of support we have received from the European Commission.

A new project is also an opportunity to look at the challenges that the community is currently facing, and frankly there is no lack of challenges: the geopolitical situation with resulting sovereignty questions, the artificial intelligence (AI) race, the reskilling and upskilling of the European workforce, the competitiveness of European industry, the ageing population, the energy transition, the war in Ukraine, growing public debt, etc. Some people are discouraged by this long list, but this is not helpful.

In my opinion, Europe is still by far the best continent to live in: it is a diverse and largely safe place with good quality of life, outstanding and affordable education, democratic governments, and excellent social protections compared to the rest of the world. Despite what some people are saying about Europe, it is not a failed continent: the European Union is the world's second largest consumer market, it is not bankrupt, it can defend itself, it is a leader in the green transition, and it is successfully competing with the rest of the world in many important sectors.

All this should make us proud and drive us to invest our energy in maintaining high standards of living in Europe, both for ourselves and for future generations. Many of the above-stated challenges can be mitigated by sustainable economic growth, based on successful innovations. There are plenty of areas in which Europe can innovate and create economic and societal value. Europe has everything it takes to lead global innovation, but success requires collective belief to start with.

My message to the HiPEAC community is that we should start spreading the word in our local communities today and kickstart the innovation movement. Innovation will be the leitmotif of HiPEAC8, with new activities including entrepreneurial excursions to thriving innovation ecosystems. We look forward to seeing what the HiPEAC community creates.

Koen De Bosschere, HiPEAC coordinator

For some years now, data has been viewed as a strategic asset – fuelling, among other things, the stratospheric ascent of generative artificial intelligence (AI). Having pioneered privacy legislation with the General Data Protection Regulation (GDPR) in 2018, the European Union (EU) has increasingly focused on how to generate value from Europe's unique data landscape, with its Data Act and Data Union strategy.

A key player in these initiatives is the Big Data Value Association (BDVA), established in 2014. HiPEAC spoke to Daniel Alonso, senior technical lead on big data and AI ecosystems at BDVA, to learn more about the strategic importance of data, how BDVA supports data-driven innovation, and what the future holds for data and AI in Europe.

'Europe must move beyond a purely technology-driven approach and ensure that data and AI generate clear economic, societal and strategic value'

BDV BIG DATA VALUE ASSOCIATION

When did the topic of data become a major concern for policy makers in the European Union, and what prompted this?

Big data emerged as a major policy topic around 2012–2013. The European Commission quickly recognized the strategic importance of data for Europe's economy and competitiveness, leading to the launch of the Big Data Value Public-Private Partnership (BDV PPP) in 2014, with a €2.5 billion investment programme that became operational in 2015. BDVA was established as the private-sector counterpart to the European Commission within this initiative. This effort was further strengthened by the European Data Strategy in 2020, which introduced the concept of European Common Data Spaces to overcome data silos, together with key data-related regulation such as the GDPR, the Data Governance Act and the Data Act. More recently, the 2025 European Data Union strategy has updated this framework with a stronger focus on data for AI and an international perspective.

"One of the major challenges in Europe is the availability of the vast amount of high-quality industrial data we need to train large AI models"

Why was BDVA created and what is its mission?

BDVA was founded in 2014 as the private counterpart to the European Commission for the BDV PPP. During the 2014–2020 period, it acted as the umbrella organization for European projects and initiatives focused on data innovation. Today, BDVA brings together a broad European ecosystem of industry, research organizations, small and medium enterprises (SMEs), and public-sector stakeholders working on data and AI. Its mission is to foster data-driven innovation in Europe by developing strategic research and innovation agendas, supporting collaboration across sectors, and providing guidance to policymakers and industry. BDVA also serves as a forum for knowledge sharing and discussion on topics related to data, AI and digital transformation at the European level.

What are BDVA's main areas of activity?

Over the 10 years since it was established, BDVA has evolved to become the centre of a vibrant ecosystem, with more than

BDVA works across several different areas to help Europe maximize data-driven innovation

250 members all around Europe. BDVA itself is a member of various European-level initiatives, including ADRA (the AI, Data and Robotics Association), EuroHPC JU (representing the perspective of AI and data, and requirements from industry) and the Computing Continuum initiative coordinated by HiPEAC (which includes around 15 associations).

With data and AI at the core of its activities, BDVA has been involved in the establishment of European Data Spaces from the very beginning, with relevant publications in the field on key topics such as metadata interoperability, data quality, privacy-enhancing technologies (PETs) and generative AI and data spaces. As part of its events, Data Week and the European Big Data Value Forum, BDVA is also leading discussions on new topics like the synthetic generation of data, data-driven compliance, AI-driven data products, and exploring new approaches to semantic interoperability.

A key focus of BDVA is creating value out of data, data sharing and data spaces. To that end, BDVA was a partner of the first phase of the Data Spaces Support Centre (DSSC) where it led the pillar on value creation and played an active role in the business, co-creation and design principles aspects, and is also a partner in the second phase, which recently started.

BDVA is also working to facilitate the establishment of AI Factories in Europe (leveraging its connection with the EuroHPC, the fact that most of the AI Factories hosting entities are BDVA members, and the knowledge of the community around the convergence between HPC, data and AI), and is also supporting the creation of Data Labs. Finally, on the standardization front, BDVA is liaising with CEN CENELEC JTC21 and JTC25, and is leading a standard on collaborative data quality.

European Big Data Value Forum 2025 - Photo: © BDVA

BDVA has been active in the AI Data Labs initiative

What are priority areas for data in Europe? What are some of the challenges which we need to overcome?

In my opinion, one of the major challenges in Europe regarding data is still the availability of (and access to) the vast amount of high-quality industrial data we need to train large AI models. Potential industrial data providers remain reluctant to share data, either because they do not know they can do so, or because of a lack of trust, or because they don't see the value of sharing their data.

Data Spaces are trying to solve this challenge in different sectors, and the new AI Data Labs have been conceived exactly to address this challenge for AI ecosystems (for now in AI Factories), but we are not yet at that point. Ultimately, the underlying challenge is not only technological, but also about value creation. Europe must move beyond a purely technology-driven approach and ensure that data and AI generate clear economic, societal and strategic value, while also strengthening European sovereignty and competitiveness.

What are the major trends you've seen in this area over the last few years?

Several important trends have shaped the data landscape in recent years, including advanced data analytics, the emergence of the data continuum, and new data architecture paradigms such as serverless computing, data virtualization and data lakehouses. These modern approaches overcome the limitations of traditional systems, particularly when dealing with the challenges of today's data landscape: vast data volumes, unstructured and multimodal formats, the growing need for real-time processing and demands for scalability and flexibility.

▶

One particularly important trend (related to the challenge of availability of high-quality data) is the growing interest in the generation of synthetic data. Synthetic data can help address the scarcity of high-quality datasets and reduce barriers linked to personal or sensitive data. At the same time, it also introduces new challenges related to quality, trustworthiness, governance and validation.

Another trend which has had a major impact on BDVA's work is the rise of AI. With data still at the core of EU innovation, we are increasingly seeing a symbiotic relationship between AI and data: AI technologies can improve data management in multiple ways, while processes have to be reviewed and redesigned to prepare data for the specific requirements for AI (see the 'AI-ready data products' paper in 'Further reading' for a framework to assess the AI readiness of different aspects of data products).

In which major areas will BDVA be active in the future? Where do you think Europe could make major wins thanks to strategic data management?

BDVA is increasingly focused on the concept of industrial AI. This approach combines data, AI and business value within a single framework to deliver tangible benefits for European industry. The objective is not only to develop advanced AI technologies, but also to ensure they create measurable value and support Europe's industrial competitiveness. A key aspect of this vision is the convergence of different European initiatives and instruments, including AI Factories, Data Spaces, AI Testing and Experimentation Facilities (AI TEFs), Data Labs and European Digital Innovation Hubs (EDIHs). Bringing these elements together can improve the efficiency, reliability and adoption of AI across European industry, while also facilitating access to high-quality industrial data.

Finally, this framework also includes some of the most prominent emerging topics in AI, such as agentic AI and physical AI, which are already being actively discussed within BDVA. Through some of our industrial members, we are involved in the IPCEI-AI project, aimed at building a sovereign AI ecosystem for European industries. We think that this initiative is fully needed and will have a strong impact on the European industrial ecosystem and the way AI is developed, deployed and adopted across industry

"Trends shaping the data landscape include advanced data analytics, the emergence of the data continuum, and new data architecture paradigms such as serverless computing, data virtualization and data lakehouses"

FURTHER READING:

Data Union factsheet
digital-strategy.ec.europa.eu/en/factpages/data-union-nutshell

BDVA publications bdva.eu/downloads

IPCEI-AI: Important Project of Common European Interest on Artificial Intelligence
bit.ly/IPCEI-AI

Data Week data-week.eu

European Big Data Value Forum european-big-data-value-forum.eu

BDVA position paper: 'Data sharing spaces and interoperability', 2024
bit.ly/BDVA_interoperability

BDVA paper: 'Elevating data quality: A paradigm shift for data spaces and AI needs', 2024
bit.ly/BDVA_data_quality

Joint BDVA and Centre of Excellence for Data Sharing & Cloud (CoE-DSC) white paper: 'Leveraging the benefits of combining data spaces and privacy enhancing technologies', 2024
bit.ly/BDVA_privacy-enhancing_technologies

Joint BDVA and Data Spaces Support Centre (DSSC) white paper: 'Generative AI and data spaces', 2024
bit.ly/BDVA_GenAI_data_spaces

BDVA workshop on new data architectures, 2025
bit.ly/BDVA_new_data_architectures

BDVA paper: 'AI-ready data products', 2025
bit.ly/BDVA_AI_data_products

HiPEAC8 kicks off with partners bringing new expertise

In May, HiPEAC8 (High Performance Edge And Cloud computing) was officially inaugurated at a kick-off meeting in Brussels. The eighth edition of the HiPEAC project brings together expertise from across the computing continuum, with Ericsson, ETP4HPC and imec joining existing partners (Ghent University, Barcelona Supercomputing Center–Centro Nacional de Computación, CEA, Inria, RWTH Aachen University, Inside Industry Association, Eclipse Foundation, Sintef and Thales).

HiPEAC8 will continue and build upon HiPEAC's well-established programme of activities, including:

- roadmapping for computing research in Europe, with the annual publication of the HiPEAC Vision;
- networking activities, including the HiPEAC conference and summer school;
- a talent hub built around the HiPEAC Jobs specialist jobs portal;
- a partnership forum bringing together representatives from relevant stakeholders in the European computing ecosystem, including joint undertakings, industry associations, and alliances, to align on high-level policy objectives.

This edition of the project has a particular focus on innovation. In addition to existing activities such as entrepreneurial courses, keynote talks and the HiPEAC Technology Transfer Awards, HiPEAC8 introduces excursions in which would-be founders travel to thriving European startup hotspots, learning lessons and building skills for their own ventures. The first excursion will take place in November; visit the HiPEAC website for updates.

hipeac.net

HiPEAC Vision 2026: Why Europe should create its own future, not imitate others

With the ripple effects from advances in artificial intelligence (AI) – mostly emanating from outside of Europe and taking place at breakneck speed – spreading relentlessly across the computing ecosystem, the HiPEAC Vision 2026 offers a calm, measured appraisal of the state of the art and offers a path forward for European computing research.

Rather than blindly copying the trajectory of other countries, it argues that Europe must chart its own course based around the 'next computing paradigm' proposed in previous editions of the HiPEAC Vision. In this vision, computing is a utility which blends seamlessly into the fabric of everyday life, offering users services on demand while eschewing the antisocial incentives of the attention economy and surveillance capitalism.

Illustrated by cartoons exclusively produced for HiPEAC by the Belgian comic artist Arnulf, this year's HiPEAC Vision has chapters dedicated to the following key topics:

- The 'next computing paradigm'
- Artificial intelligence
- New hardware
- Tools
- Cybersecurity
- Open source
- Sustainability
- State of the European Union

Each chapter has its own recommendations, while a consolidated list of recommendations is also available, grouped into technological, standardization, methodological and policy recommendations. This year, HiPEAC has also developed a tool which uses the HiPEAC Vision as the exclusive basis for returning answers via AI chatbots, which can be connected to your chatbot of choice via HiPEAC's Model Context Protocol (MCP) server. This is complemented by a dedicated tool to explore the text from different angles, such as the policy, industry, investment or educational perspective.

FURTHER READING:

hipeac.net/vision
HiPEAC MCP server hipeac.net/mcp
Dedicated tool to query the HiPEAC Vision ask.hipeac.net
CONNECT University presentation at the European Commission
bit.ly/HiPEAC_Vision_2026_CONNECT_University

Sage workshop at DISCOVER-US

This is a summary of a detailed report by Pete Beckman, the director of the Sage Grande Testbed, which is building a cutting-edge artificial intelligence (AI) cyberinfrastructure to support advanced AI research. To read the full report, visit bit.ly/SAGE_DISCOVER-US_2026

On 4-6 May 2026, around forty researchers from across the United States and Europe gathered in Barcelona for a workshop on AI and edge computing for scientific research. The meeting was hosted at Universitat Politècnica de Catalunya (UPC), and was supported by the DISCOVER-US project, which is funded by the European Union (EU). The workshop was organized by Pete Beckman (Northwestern University), Nicola Ferrier (Argonne National Laboratory / Northwestern), and Rosa M. Badia (Barcelona Supercomputing Center–Centro Nacional de Computación), and was kicked off with a welcome talk by Koen De Bosschere (Ghent University) on behalf of DISCOVER-US.

The participant list was deliberately broad. Ecologists and biodiversity researchers sat alongside computer scientists and AI / orchestration specialists, hazards researchers, and deployment leaders. The participants were split roughly evenly between the United States and Europe.

Format and flow

The three-day programme included short framing talks to set the intellectual stage, lightning talks from domain experts, and then structured breakout discussions where small groups wrestled with a small number of carefully worded questions. The four thematic sessions worked through the problem space from complementary angles.

Session 1 (Ecology and Biodiversity) was framed by Anne Fouilloux (LifeWatch) and Christine Kreuder Johnson (UC Davis), with lightning talks from Joana Colomer (LTER Spain/ Minuartia), Nicolas Moity (Charles Darwin Foundation, Galápagos), Sydne Record (University of Maine), and Rachel Poretsky (UIC).

Session 2 (Computer Science – AI and Orchestration) was framed by Nicola Ferrier on AI and Francesc Lordan (BSC) on distributed computing, with lightning talks from Delia Velasco (CSIC), Elizabeth Tranel (Iowa State), Matthew Thompson (Ohio State), and Germán Moltó (UPV).

Session 3 (Hazards – Wildfire and Earthquake) was framed by Josep de la Puente (BSC), with lightning talks from Daniel Balouek (INRIA) and Leonardo Mingari (GEO3BCN).

Session 4 (Deployments and Networks) was framed by Chris Florian (NEON), with lightning talks from Sebastien Biraud (LBNL) and Gary Bucciarelli (UC Davis). The Wednesday wrap-up added perspectives from Manish Parashar (Utah) on the National Discovery Cloud, Julia McKenna (Northwestern) on the STRONG Manoomin collective, Kyle Kovach (Wisconsin) on hyperspectral sensing, from Josep de la Puente (BSC) on Distributed acoustic sensing, Delia Velasco on the Doñana Park infrastructure, and others.

Cross-cutting scientific questions and novel insights

Across every breakout, participants returned to what makes a scientific question transformative and only addressable through AI at the edge: real-time, continuous, multi-modal observation; the ability to trigger sampling on events rather than schedules; the move from correlational to causal inference; and the capacity to act, not merely record. Ecologists framed this as detecting short disturbances, species interactions, or shifting baselines that simply do not survive batch upload; computer scientists framed it as autonomous discovery, in-the-wild fine-tuning, and adaptive sampling.

The vision of the agentic edge provoked intense debate. Participants noted the enticing potential of deploying edge agents, for example orchestrating sensing, negotiating interpretations with high-performance computing (HPC) simulations, and aggregating local knowledge into shared understanding without moving raw data. However, hard problems remain: trust, verifiability, uncertainty quantification, safety certification, prompt injection, credential delegation, and the self-reinforcing biases that emerge when agents learn from each other.

On the last day, participants laid plans for ongoing interdisciplinacy collaborations in this fascinating and fast-moving area.

Last phase of DISCOVER-US to apply decentralized intelligence knowledge to EU priorities

In January 2024, the DISCOVER-US project set out to create new research alliances in distributed computing and swarm intelligence between researchers in the European Union (EU) and the United States (US), with research exchanges and vision building integral to the project's goals.

Due to changes in policy priorities in the US, the outline of the project has shifted since its kick-off at HiPEAC 2024, with the initial plan being adapted into a tripartite structure that allows the knowledge gained to be fully leveraged for key EU priorities.

The first phase of DISCOVER-US involved building a transatlantic community of researchers working on topics related to distributed computing and swarm intelligence, with researchers undertaking exchanges at top universities and research institutions in the US. These exchanges allowed EU researchers to delve into research lines from multi-agent systems to interoperability between large infrastructures and associated edge devices. They have resulted in paper publications, book editions, and plans for further joint working.

DISCOVER-US also organized vision workshops and released a joint vision on distributed computing and swarm intelligence, available for download from the project's website. Meanwhile, the project's launch the project's technical webinar series, delivered by world-class researchers, has provided perspectives on key topics in the field from both sides of the Atlantic.

In the second phase, the project focused on platforms and frameworks of common interest – such as the Sage Grande, FABRIC and Chameleon testbeds – considering industry perspectives and investigating how to scale up these platforms. Courses in the DISCOVER-US track at the HiPEAC summer school, ACACES, in 2024 and 2025 included detailed explorations of Sage and Chameleon, while a number of webinars were also dedicated to these and other platforms and frameworks.

The project is now entering a third phase, which will apply the knowledge gained during the first and second phases to European priority areas, including the energy grid, agriculture, mobility and more. To this end, two workshops have been held on unmanned aerial vehicles (UAVs), allowing representatives from European industry and academia to advance their knowledge of next-generation drone technologies. The project will culminate in a vision workshop in November 2026, which will help European policy makers shape future research in this fast-moving area.

Presentations from the DISCOVER-US session at HiPEAC 2026

Further information:

discover-us.eu

DISCOVER-US videos on HiPEAC TV

bit.ly/HiPEACTV_DISCOVER-US

Frontiers in Complex Systems edition showcases DISCOVER-US research

Edited by DISCOVER-US members Ovidiu Vermesan (SINTEF), Chandra Krintz (University of California, Santa Barbara) and Rosa M. Badia (Barcelona Supercomputing Center–Centro Nacional de Supercomputación), a special edition of the journal *Frontiers in Complex Systems* focuses on intelligent self-organized, dynamic, and adaptive complex distributed systems.

The journal features a number of papers authored by members of the DISCOVER-US community, on topics from workload balancing to code optimization with neural networks. Papers are open access and may be found on the Frontiers in Complex Systems website.

FURTHER READING:

bit.ly_DISCOVER-US_Complex_Systems_CfP

Semidynamics and SiPearl announce strategic cooperation

Semidynamics, an advanced computing company developing memory-centric artificial intelligence (AI) infrastructure for large-scale inference, and SiPearl, the European fabless designer of high-performance energy-efficient central processing units (CPUs) for sovereign supercomputing, AI and data centres, have entered into a strategic partnership to develop a European rack-scale AI compute platform dedicated to large-scale AI inference in the cloud.

The official announcement noted that the companies share a common goal: to offer a sovereign high performance energy efficient compute solution capable of supporting major European initiatives, both public and private, including AI Factory and Giga Factory programmes. Semidynamics and SiPearl will coordinate their marketing and sales efforts to jointly pursue European procurement opportunities.

According to the announcement, this joint platform will bring together core European technologies. SiPearl's Arm® based CPU will provide general-purpose compute, orchestration and data-plane hosting, while Semidynamics' RISC-V-based graphics processing unit (GPU) / AI inference application-specific integrated circuit (ASIC) will act as the main acceleration engine for AI inference workloads and enable future performance scaling. The companies expect to offer a rack-scale system delivering the density expected from leading global AI platforms. The rack design will be based on Open Compute Project (OCP) standards, supporting interoperability and alignment with established cloud and data centre infrastructure practices.

'SiPearl is thrilled to see the impact of years of work in the European Processor Initiative and the EU sovereign ecosystem come to fruition with this platform. It demonstrates the systematic progress that Semidynamics and SiPearl have made individually and collectively and will showcase the best of both companies, CPU and accelerator,' said Philippe Notton, SiPearl's chief executive and founder, and a HiPEAC member.

'We are delighted to work with SiPearl and to offer a European CPU as part of our AI inference platform,' said Roger Espasa, the chief executive of Semidynamics, also a HiPEAC member. 'Combining SiPearl's high-performance CPU with Semidynamics' RISC-V-based GPU/AI inference technology gives Europe a credible path towards sovereign, rack-scale AI infrastructure built around European-controlled compute.'

bit.ly/Semidynamics_SiPearl_2026

Axelera AI raises over $250 million in latest funding round

Eindhoven-headquartered Axelera AI, which provides artificial intelligence (AI) acceleration hardware, has announced its latest funding round led by Innovation Industries. This brings the total amount of funding raised by Axelera AI to $450 million in equity, grants and venture debt since the company's incorporation in July 2021.

'Data centres are hitting power and cooling limits, and as analytics move closer to where data is being created, edge AI solutions must operate within strict energy and bandwidth constraints,' said HiPEAC member Fabrizio Del Maffeo, chief executive and co-founder of Axelera AI. 'Our edge-first approach … [is] about making AI deployment economically viable at scale for real-world applications while protecting data and privacy by processing customer information locally.'

Axelera AI's ecosystem, the Partner Accelerator Network, brings together software vendors, model makers, system integrators, solution providers,

and technology partners to accelerate customer deployment and reduce time-to-production.

According to the company's announcement, this new funding will accelerate Axelera AI's manufacturing scale, expand its customer success organization and Partner Accelerator Network, and continue advancing its software tools and software development kit (SDK).

BSC and UPC announce new spinoff, Safe and Secure Technologies

Barcelona Supercomputing Center – Centro Nacional de Supercomputación (BSC-CNS) and the Universitat Politècnica de Catalunya - BarcelonaTech (UPC) have announced a new spinoff company, Safe and Secure Technologies. The company brings together a set of hardware technologies specifically designed to allow medium- and high-performance processors to be used in critical applications where design and validation processes are particularly demanding and often regulated by standards that guarantee functional safety and cybersecurity.

'The hardware we developed can only fail under very exceptional conditions, and when it does, it detects it and interrupts the process in a controlled manner before giving erroneous instructions. In an air-traffic control system, for example, this can be crucial to saving lives,' explained HiPEAC member Jaume Abella, co-founder of Safe and Secure Technologies S.L. and co-director of the High Performance Embedded Systems (HPES) laboratory at BSC.

The company's flagship technology, Safety Island, combines technologies developed over more than 10 years by researchers from BSC and UPC, within the context of EU-funded projects such as De-RISC, SELENE, ISOLDE, and FRACTAL. By focusing on chip design under the RISC-V open-source architecture standard, Safe and Secure Technologies supports progress toward European technological sovereignty by eliminating dependencies and licences from external multinationals, which is crucial for the critical sectors where the company operates.

bit.ly/Safe_and_Secure_Technologies_news

Mosaic SoC announces $3.8 million pre-seed round

In April, ETH Zürich spinoff Mosaic SoC announced that it had successfully raised $3.8 million in a pre-seed round led by Founderful, with participation from Kick Foundation.

Founded by Alfio Di Mauro and Moritz Scherer, who studied in the Digital Circuits and Systems Group of HiPEAC members Luca Benini and Frank K. Gurkaynak in the Integrated Systems Laboratory at ETH Zürich, Mosaic SoC designs specialized chips for real-time perception on a low power budget. According to the company's announcement, this enables spatial intelligence in wearable form factors such as augmented reality (AR) glasses.

Photo credit: Daniel Kunz

'Spatial intelligence shouldn't require an application-class processor and a GPU,' said Alfio Di Mauro, chief executive and co-founder of Mosaic SoC. 'We built Mosaic SoC to deliver real-time perception at a fraction of the energy, so battery-powered devices can understand their environment without compromising form factor.'

According to Mosaic SoC, their chip lets a device build a local map of its surroundings and the objects within them, enabling features like recalling where an item was last seen or generating a floorplan on the fly. In smartphones, Mosaic SoC can act as a co-processor for the front camera, running always-on tracking and classification at a fraction of the power. That means a device can trigger recording only when a specific event occurs or a certain object appears, delivering continuous awareness without draining the battery.

FURTHER INFORMATION:

mosaic-soc.com

Moritz Scherer's thesis featured in *HiPEACinfo 73*

bit.ly/HiPEACinfo73_Moritz_Scherer

NOfire AI raises $2.5 million in seed funding for software reliability platform

In May, NOFire AI raised $2.5 million in seed funding in a round led by Marathon Venture Capital. The company builds tools that monitor and resolve software incidents, helping engineering teams catch problems before they escalate into outages – a challenge growing ever more pressing as more production code is generated by artificial intelligence (AI).

NOfire AI's co-founder and chief scientist is HiPEAC member Anastassios Nanos, director of Nubificus Ltd. He joins the company's chief executive Spiros Economakis, chief technology officer Panagiotis Moustafellos and chief operating officer Antonios Chalkiopoulos.

Posting on LinkedIn, Anastassios said: '[I'm] so excited to join forces with Panagiotis Moustafellos and Spiros Economakis at NOFire AI to help define and deploy the next-generation runtime for agentic workloads, bringing what we've built at Nubificus Ltd to bear on attestation, isolation, and policy enforcement for software that acts on its own. ... The future of autonomous software is only as trustworthy as the ground it runs on.'

In a blog post announcing the seed round, Marathon Venture Capital partner Panos Papadopoulos noted: 'While the velocity of code creation has scaled exponentially, the engineering lifecycle dedicated to running and maintaining production systems has remained stubbornly human-dependent. Site Reliability Engineering (SRE)—the critical function responsible for keeping digital platforms available, fast, and healthy—is hitting a hard wall. [...] We need an AI SRE layer capable of cross-domain reasoning and autonomous remediation at machine speed.'

nofire.ai

Vertical Compute raises € 57 million to commercialize memory-on-logic chiplets

Leuven-headquartered chip company Vertical Compute recently raised € 37 million in a seed round led by Quantonation, in addition to € 20 million raised earlier in the year. Vertical Compute has developed a 3D chip architecture that integrates memory directly above compute logic within a single wafer manufacturing process. According to the company, this approach reduces data movement from centimetres to nanometres, thereby minimizing latency and improving energy efficiency. This latest round of investment will be used for commercial chiplet deployment for future artificial intelligence (AI) platforms, with chiplets designed to complement existing processors and accelerators.

verticalcompute.com

Fractile announces $ 220 million funding for AI inference chips and systems

UK-based Fractile announced in May that it had raised $220 million in Series B funding for its artificial intelligence (AI) inference processors, in a round led by Accel, taking it to 'unicorn' status. Fractile is commercializing in-memory chips which it says will significantly reduce the time required for AI inference workloads.

Announcing a previous investment round on LinkedIn, Pat Gelsinger, former chief executive of Intel and one of the company's angel investors, said: 'Being able to run any given model orders of magnitude faster, at a fraction of the cost and maybe most importantly at dramatically lower power envelop [sic] provides a performance leap equivalent to years of lead on model development.'

fractile.ai
bit.ly/Pat_Gelsinger_LI_Fractile

Funding bonanza for European quantum computing companies

Recent months have seen a flurry of announcements about successful funding rounds for European quantum computing companies. Among the companies attracting large rounds of investment are the following.

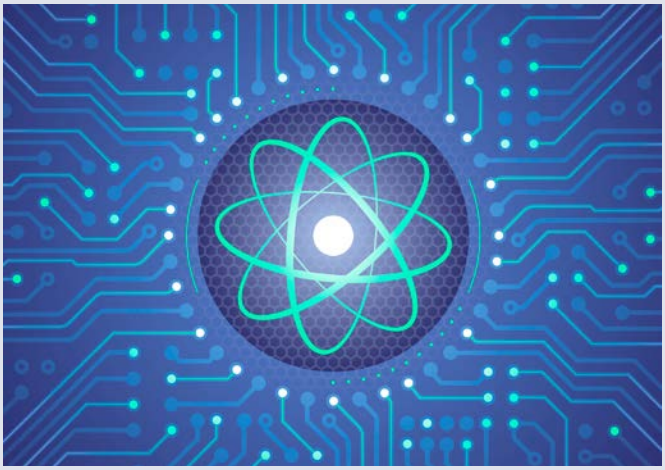


Image credit: Sergey Tarasov | stock.adobe.com

Oxford Quantum Circuits (UK): £ 260 million series C led by Bullhound Capital, June 2026

Founded in 2017, Oxford Quantum Circuits is a spinout from Oxford University which develops and operates superconducting quantum computers for real-world applications, headquartered in Reading, UK. The current round of funding will be used to support the company's global expansion, including scaling quantum infrastructure in key markets and advancing the development of their systems.

oqc.tech

QuantWare (Netherlands): $ 178 million series B co-led by Intel Capital and FORWARD.one, May 2026

A spinout of Delft-based QuTech research institute, QuantWare focuses on superconducting circuits and designs quantum processing units (QPUs). The funding round was made public after the company, which aims to be the 'Intel of quantum computing', announced VIO-40K™, a quantum processor architecture for 10,000 qubits designed as an open platform that can scale third-party qubit chiplets and designs. The company is building KiloFab, which it describes as 'the world's largest dedicated quantum open architecture fab, increasing the company's production capacity by 20x to meet strong global customer demand'.

quantware.com

Quobly (France): € 115 million series A led by Bpifrance, SEALSQ and STMicroelectronics, June 2026

Headquartered in Grenoble, Quobly develops silicon-based quantum chips using established semiconductor manufacturing processes. The funding will be used for continued research and development (R+D), industrialization efforts and international commercial expansion. Quobly, which was spun off from CEA Leti in 2022, plans to deploy its first commercial quantum computer through the cloud by the end of 2026 under its Alloy product line.

quobly.io

eleQtron (Germany): € 57 million series A led by Schwarz Digits, May 2026

eleQtron, a developer of trapped-ion quantum computers, was founded in 2020 as a spinoff of the University of Siegen. According to the company's announcement, this capital will be used to build scalable production capacity, expand cloud-based access to eleQtron's systems, and further advance development of its hardware platform. In parallel, eleQtron is advancing its proprietary MAGIC technology (Magnetic Gradient Induced Coupling). MAGIC uses radio frequency and microwave technology to control qubits, which eleQtron says is both more precise and more robust than laser control of qubits.

eleqtron.com

IQM (Finland): € 50 million financing package from funds and accounts managed by BlackRock, following $ 300 million series B led by Ten Eleven Ventures

Headquartered in Espoo, IQM was founded in 2018 as a spinoff from Aalto University and VTT Technical Research Centre of Finland. The company, which manufactures superconducting quantum computers for research institutes, universities, and high-performance computing (HPC) centres, announced in September 2025 that it had raised $ 320 million (€ 275 million) to expand its commercial presence and scale its data-centre infrastructure and assembly lines globally. In March, IQM announced that it had secured an additional € 50 million financing package to support acceleration of its technology roadmap, fuel R+D, support entry into additional markets, and advance the company's leadership in quantum computing as it prepares for public listing.

iqm.tech

Groove Quantum (Netherlands): €16 million seed round led by Innovation Industries and 55 North

A spinoff from Delft University of Technology (TU Delft) and QuTech, Groove Quantum is developing quantum computers based on germanium semiconductors. The company states that its germanium qubits are high quality with a small footprint and integrate into existing semiconductor manufacturing processes, all accelerating scaling. The funding will be used to scale qubit counts, grow the team, and initiate industrial-grade manufacturing at leading semiconductor foundries.

groovequantum.com

Recent EIC results – and what they mean for the European computing ecosystem

Paul Pietrangelo, Lira

2025 was a good year for computing researchers applying to Europe's flagship innovation programme, the European Innovation Council (EIC). With the main funding schemes remaining largely the same in 2026, there are still opportunities to get your deeptech venture funded, joining several others in the HiPEAC community who have been successful in this highly competitive space.

EIC Pathfinder

The **EIC Pathfinder** funds research of radical and visionary ideas (technology readiness levels (TRLs) 1 to 4). Although high risk, these projects should nevertheless have the potential for high reward, such as a significant technological advancement that could transform industries or create entirely new markets. The programme offers single applicants or consortia grants of up to €4 million to achieve proof of principle and validate the scientific basis of breakthrough technologies. As an example, the CERBERUS project on p.20 of this magazine was funded under the 2024 EIC Pathfinder Challenge Call.

In 2025, the EIC Pathfinder Open Call received 2,087 proposals, from which 44 projects were chosen, sharing a budget of €140 million. Among them were six advanced computing or related projects, including:

- **SuperICQ**: Superconducting Integrated Circuits for Scalable Quantum Systems
- **NEQIOS**: Neuromorphic Quantum-Inspired Optimization and Simulation
- **CEREBRIS**: Multi AI-Agents to Revolutionise the Management of Neurological Disease

There were 667 proposals submitted to the EIC Pathfinder Challenges Call, of which 30 projects were chosen, receiving €118 million. For computing researchers, seven projects were funded under the 'Generative-AI based Agents to Revolutionize Medical Diagnosis and Treatment of Cancer' call and eight projects were funded under the 'Towards autonomous robot collectives delivering collaborative tasks in dynamic unstructured construction environments' call.

In 2026, the EIC Pathfinder has a budget of € 286 million, with €166 million allocated to the open call where breakthrough ideas from any field of research can apply, and €120 million allocated to three challenge calls, each with €40 million. The open call, which closed in May, received 2,013 proposals with 4,633 unique beneficiary institutions spread across 76 countries.

The deadline for the challenge calls is 28 October. Among the calls is one titled 'DeepRAP: Deep Reasoning, Abstraction & Planning towards trustworthy Cognitive AI Systems'. The goal of this challenge is to move beyond the current state of the art in traditional, deep-learning and reinforcement-learning paradigms, with a focus on significantly improving the reasoning, abstraction, and planning (RAP) capabilities of artificial intelligence (AI) systems.

EIC Transition

The **EIC Transition** funds the maturation and validation of a novel technology and its accompanying business plan from the lab to the relevant application (TRLs 4 to 5/6). Grants of up to €2.5 million are offered to single applications or consortia.

In 2025, 611 proposals were submitted to the EIC Transition call, with 40 projects being selected, sharing a total budget of €96 million. Impressively, 25% of the projects selected were from advanced computing or related fields, including:

- **BELFORT ASIC**: Pioneering Hardware Acceleration for Computing on Encrypted Data
  Submitted by HiPEAC Technology Transfer Award 2025 winner Belfort
- **TRU-STACK**: Transparent Resource Utilisation for Scientific and Enterprise Workloads
  Submitted by the System Tools and Advanced Runtimes (STAR) group at Barcelona Supercomputing Center-Centro Nacional de Supercomputación, led by HiPEAC member Vicenç Beltran
- **RE-CLIQS**: REplaceable Chiplet LInks for Quantum Scaling
  Submitted by QuantWare, a Netherlands-based company which has just had a successful $178 million Series B funding round (see p.13)

In 2026, the EIC Transition has a budget of €100 million. The deadline for the call is 16 September.

EIC Accelerator

The **EIC Accelerator** funds single start-ups and small / medium enterprises (SMEs) that are developing high-impact, deep-tech solutions that can create new markets or disrupt existing ones. It offers these companies grants of up to €2.5 million and equity investments of up to €10 million to fund the later stages of technology development as well as scale up (TRLs 6 to 9).

In 2025, out of 1,882 full applications, there were 101 EIC Accelerator winners. Among the winning companies, 12% were digital, 7% quantum, and 3% semiconductor. Notable advanced computing winners include:

- **roofline AI**, an Aachen-based AI compiler company co-founded by HiPEAC members Jan Moritz Joseph and Rainer Leupers
- **SpiNNcloud**, which used EIC Transition funding to take SpiNNaker technology, developed by HiPEAC member Steve Furber and colleagues, to the edge
- **Magics**, a company based in Belgium which develops radiation-tolerant, high-reliability integrated circuits
- **Quantum Dice**, an Oxford-based company developing quantum photonics and probabilistic computing

In 2026, the EIC Accelerator has a budget of €634 million, with €414 million allocated to the open call for companies from any industry, and €220 million allocated to five challenge calls, although this year there is no challenge call focused on advanced computing research. In 2026, deadlines for full applications are every two months.

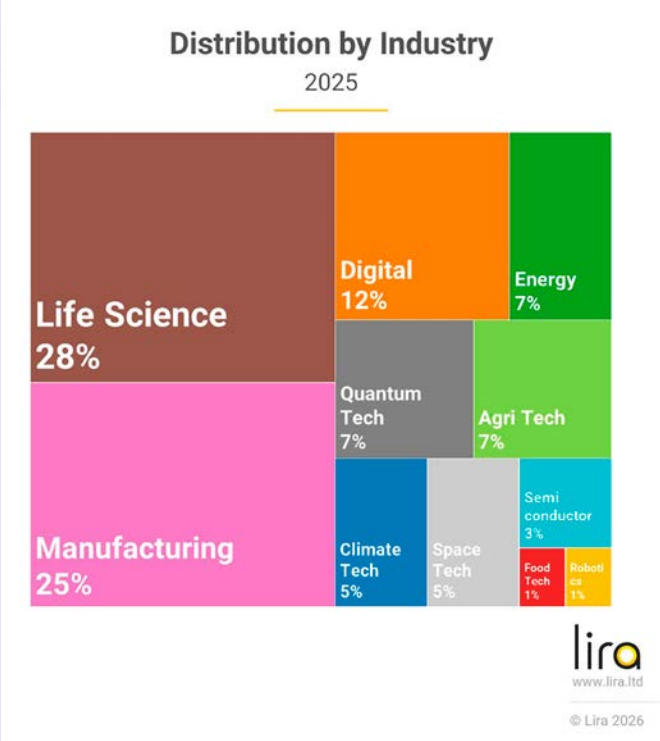


Image credit: Lira

EIC STEP Scale Up

Lastly, 2025 was the first year of a new funding programme: **EIC Strategic Technologies for Europe Platform (STEP) Scale Up**. This supports the development and manufacturing of critical technologies, including digital technologies and deep-tech innovations. It does so by offering scale-ups in line with the EU's key strategic priorities equity funding up to €30 million.

In 2025, 104 companies applied to the programme, of which 19 were selected for funding, 12 of which address digital technologies and deep-tech innovation, including Axelera AI – which is led by HiPEAC members Fabrizion Del Maffeo and Evangelos Eleftheriou and advised by HiPEAC members Luca Benini, Marian Verhelst and Torsten Hoefler – Multiverse Computing, and Pasqal.

In 2026, the EIC STEP Scale Up programme has a budget of €300 million and there is a specific focus on semiconductor and quantum technologies. In 2026, there is a deadline every quarter; in May, the first results were announced, with two quantum computing companies (Alice & Bob and QuantWare) among the eight winners.

Further information about these calls and data analysis:

Lira website
lira.ltd

Lira on YouTube
youtube.com/@TheDataGuy_Lira

Paul Pietrangelo delivered the entrepreneurial course at the HIPEAC summer school, ACACES, in 2020, 2021 and 2024.

"Impressively, 25% of the [EIC Transition] projects selected were from advanced computing or related fields"

TASKING announces toolchain enhancements for agentic AI workflows optimizing the development of functionally safe and secure systems

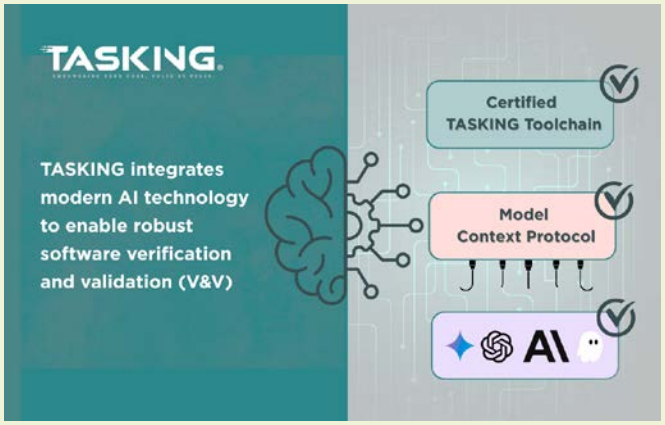
Earlier this year, TASKING, a major provider of high-performance embedded software development tools, announced enhancements to its toolchain that enable seamless integration of artificial intelligence (AI) in the software development and verification workflows. According to the company's announcement, these new capabilities accelerate the design and increase the performance of functionally safe and secure embedded real-time applications within automotive, aerospace & defence, industrial, and robotics while enabling original equipment manufacturers (OEMs) to verify and validate (V&V) systems using agentic AI workflows.

While traditionally workflows, processes, and tests have been designed by hand, developers can now use large language models (LLMs) to direct external AI agents and tools to automate many repetitive manual design, debug, and testing tasks. For example, AI-assisted workflows can implement comprehensive testing in earlier design stages to identify and resolve a wide range of issues sooner. According to TASKING, the result is faster time to market with less chance of human error, greater overall system reliability, and lower development investment.

'AI enables today's developers to be both more productive and efficient while at the same time delivering higher software performance and quality,' said Christoph Herzog, co-CEO of TASKING. 'By taking over tedious and time-consuming tasks, AI-assisted tools can free up individuals to focus on value-added design. With the AI capabilities of the TASKING tools, development teams can now develop, verify, and validate complex systems faster, with less risk, and at a lower cost.'

AI is playing an increasingly significant role in the design, debug, and test of real-time embedded applications. However, TASKING notes, as AI is often implemented as a probabilistic tool, it can produce different results each time it is used. For systems that need to verify and validate deterministic behaviour, AI must be carefully introduced to the workflow in a manner that enables systems to adhere to strict industry standards. For the foreseeable future, humans will remain in the development loop, so making them more productive and efficient in these processes is a differentiator to organizations willing to leverage AI.

> *"Development teams can now develop, verify, and validate complex systems faster, with less risk, and at a lower cost"*



The TASKING toolchain has been designed with a foundation that enables OEMs to develop functionally safe and secure systems. Modern AI capabilities are supported within the toolchain using the Model Context Protocol (MCP), an open-source standard that allows AI agents to securely interact within the development tools and access data required to achieve associated tasks. In this way, developers can use an LLM to control AI agents that direct and automate key aspects of the development lifecycle.

'With the TASKING toolchain, AI can become an integral part of functionally safe and secure workflows,' said Janez Ulcakar, director of research & development, TASKING. 'Workflows also become more flexible and agile, enabling developers to continuously optimize and enhance code with productivity enhanced by AI assistance. This gives OEMs the competitive edge of not just improving code design but of optimizing their entire software development lifecycle.'

bit.ly/TASKING_AI_agents_workflow
For further information on the general workflow and how AI agents work within TASKING's tools, see this video: 'Powering Agentic AI Workflows in Embedded Development with TASKING'

https://youtu.be/Yd9M8sBg2Ew

EXA4MIND AI Inference Service Solution

Tomáš Martinovič, IT4Innovations

High-performance computing (HPC) centres increasingly face the need to support artificial intelligence (AI) inference services, even though they are traditionally designed for batch workloads governed by strict allocation rules. Their infrastructure is not optimized for dynamically scalable, on-demand services such as large language model (LLM) queries. Nevertheless, providing AI inference capabilities can be valuable in specific cases, particularly for agentic systems and for the use of LLMs within controlled and trusted environments where security and data privacy are essential.

The EXA4MIND AI Inference Service Solution bridges the gap between cloud-native artificial intelligence and high-performance computing by providing a persistent, scalable, and user-friendly large language model inference service running on HPC resources accelerated by graphics processing units (GPUs). The service offers an OpenAI-compatible REST application programming interface (API), enabling smooth integration with external applications. Under the hood, the service relies on HPC job orchestration through HEAppE (HPC-as-a-Service Middleware) and uses job pre-allocation to provision new compute resources before wall-time expiration, ensuring continuous service availability. The system can deploy the vLLM engine, or the Triton inference server.

We tested the server on the LUMI supercomputer using the HuggingFace Inference Benchmarker, running Qwen/Qwen3-4B-Instruct-2507 in two configurations with increasing queries per second (QPS): vLLM on a bare-metal compute node and vLLM on a compute node accessed through the AI inference service. The EXA4MIND Inference Service demonstrates strong scalability, maintaining a throughput within 2% of bare-metal performance. Although HTTP overhead results in a higher time to first token (TTFT), the service successfully handled a sustained load of 19 requests per second over 120 seconds, completing 2,256 requests compared to 2,260 on bare metal.

By bringing together the resource predictability and high utilization of traditional HPC environments with the agility and user-friendly access characteristic of cloud services, this architecture eliminates operational barriers for AI users while still meeting the strict scheduling and governance requirements of supercomputing facilities. It enables HPC centres to deliver reliable, production-grade AI inference at scale, positioning them at the forefront of the rapidly growing convergence of cloud, AI, and HPC technologies.

FURTHER READING:

EXAMIND AI Inference Service inference.exa4mind.eu

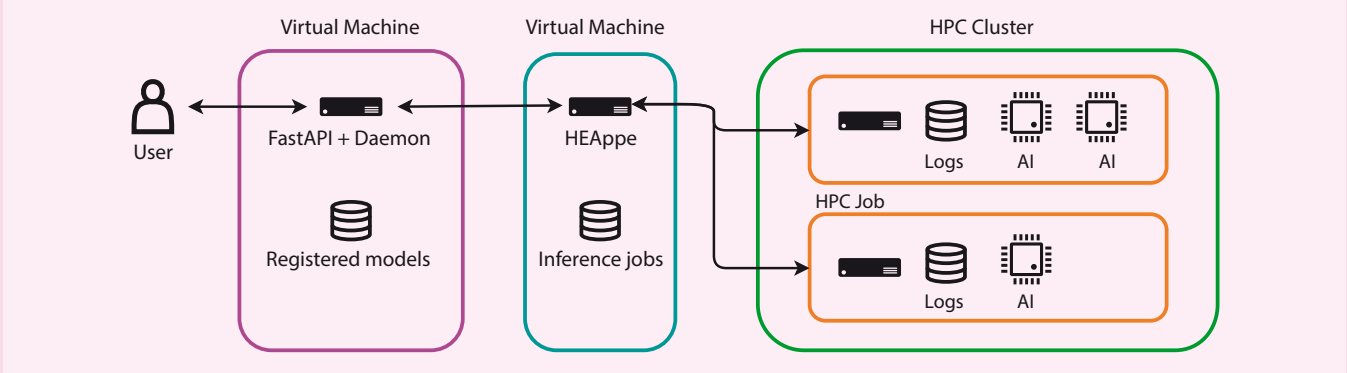
EXAMIND project exa4mind.eu

HEAppE (HPC-as-a-Service Middleware) heappe.eu

HuggingFace Inference Benchmarker
github.com/huggingface/inference-benchmarker

This research was supported by the EXA4MIND project, funded by the European Union's Horizon Europe Research and Innovation Programme, under grant agreement no. 101092944. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Commission.

Funded by the European Union



EXA4MIND AI Inference Service Architecture

CEA demonstrates first dynamically routed electro-optical router for photonic interposers

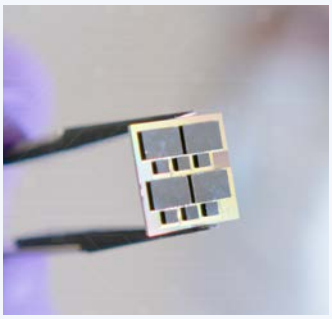


Photo credit: KAM Productions

At ISSCC 2026, the IEEE International Solid-State Circuits Conference (ISSCC), CEA researchers presented the first electro-optical router with dynamic, frame-level optical routing integrated with complementary metal-oxide semiconductor (CMOS) control logic. According to CEA's announcement, this marks a major step toward practical optical networking inside advanced chiplet-based packages.

Current optical interconnects are largely limited to static, point-to-point links, with initialization and training times ranging from microseconds to milliseconds. While suitable for board-level or rack-scale communication, those latencies prevent optical links from being used as a true networking fabric inside multi-die packages.

The router addresses this gap by integrating optical switching, routing control, serializer / deserializer (SerDes), and clocking logic directly with silicon photonics. The result is a dynamically routed optical interconnect that operates at nanosecond timescales, enabling optical communication across centimetre-scale interposers with responsiveness previously limited to short electrical links.

The prototype is fabricated in a 28 nm CMOS process and integrated on a photonic interposer. Compact analogue drivers, combined with standard-cell-based SerDes and clocking circuits, enable dense integration of optical endpoints close to compute and memory resources.

While the architectural target includes central processing units (CPUs), graphics processing units (GPUs), and high-bandwidth memory (HBM) in large 2.5D and 3D packages, the current chip serves as a proof of concept, demonstrated on a small-scale multi-die system derived from CEA-Leti's earlier INTACT active interposer architecture (ISSCC 2020).

'As chiplet systems continue to grow in scale and complexity, the ability to move data efficiently across the entire package becomes essential,' said CEA-List's Yvain Thonnart, lead author of the paper. 'Our goal was to demonstrate that photonic links can provide that reach without sacrificing the flexibility designers expect from modern interconnects. This router is a step toward practical, dynamically routed optical networks that fit within standard CMOS design flows and real product constraints.'

FURTHER INFORMATION:

Y. Thonnart et al., 'A 3.19pJ/b Electro-Optical Router with 18ns Setup Frame-Level Routing and 1-to-6 Wavelength-Flexible Link Capacity for Photonic Interposers', 2026 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 2026, pp. 1-3, doi: 10.1109/ISSCC49663.2026.11409067.

Call for posters: REACH 2026

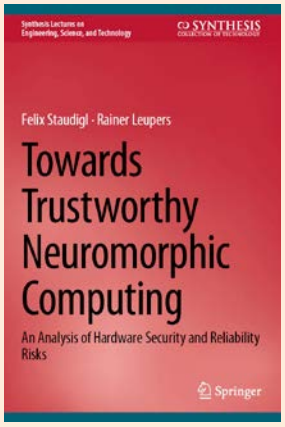
REACH 2026 (Reach Emerging Architectures in Computing Horizons) will take place on 9-10 November 2026 at the Institution of Engineering and Technology in London. The call for posters is currently open, with a deadline of 31 July.

Now in its third year, the Institution of Engineering and Technology's computer architecture conference REACH brings industrial and academic researchers together to shape the future of computer science and engineering.

Several HiPEAC members, including HiPEAC founding member Olivier Temam (Google DeepMind), Marian Verhelst (KU Leuven), Babak Falsafi (EPFL) and Kerstin Eder (University of Bristol) are scheduled to speak at the 2026 edition, along with senior representatives from Arm, Meta, GraphCore and Quantinuum.

For more information, and details on how to submit, visit the REACH website.
bit.ly/REACH26_HiPEAC

Navigating the challenges of computing in memory



Felix Staudigl and Rainer Leupers, RWTH Aachen University

The von Neumann bottleneck – the persistent lag caused by separating memory and processing – has long hampered computing performance. While computing in memory (CIM) promises to bridge this gap, it also presents new reliability challenges.

While memristors offer the non-volatility and high density needed for brain-inspired computing, they bring a specific set of physical realities: device variability, ageing effects, and, consequently, new hardware-level security risks. Our book, titled *Towards trustworthy neuromorphic computing: An analysis of hardware security and reliability risks*, provides a technical deep dive into these hurdles, offering a structured look at how to build reliable and secure memristive-based CIM systems.

Key highlights of this volume include:

- **A multi-level fault injection platform**
  X-Fault and **FLIM** represent a fault injection framework designed to simulate reliability, from individual memristor gates all the way up to full-fledged workloads.
- **The NeuroHammer attack**
  A detailed analysis of a novel security vulnerability in resistive random-access memory (ReRAM). Similar to dynamic access memory (DRAM)'s Rowhammer, this attack exploits thermal crosstalk to induce bit-flips, potentially compromising sensitive data like RSA keys.
- **The NeuroBreakoutBoard (NBB)**
  A look at the design and implementation of a versatile platform for characterizing memristive devices at the crossbar and operational levels.

By examining the intersection of **analogue CIM** and **logic in memory (LIM)**, the text moves beyond theoretical benefits to address the practical concerns of hardware-enabled attacks and device-level instability. For researchers and engineers working on the next generation of integrated circuits, this book serves as a roadmap for understanding – and mitigating – the risks inherent in memristor-based systems.

bit.ly/trustworthy_neuromorphic

New book bridges embedded testing fundamentals and AI-based techniques

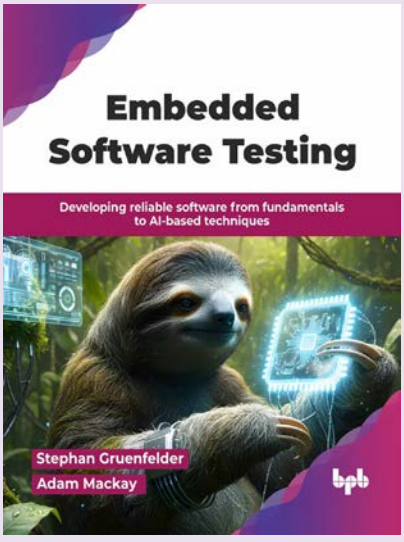
HiPEAC member Adam Mackay (Head of AI, QA-Systems) and Dr Stephan Grünfelder (Senior Engineer, Heicon & Riedel Communications) have published *Embedded Software Testing: Developing Reliable Software from Fundamentals to AI-based Techniques* (BPB Publications, 2026). The book was launched at Embedded World in Nuremberg in March 2026.

The book covers the full spectrum of embedded software testing practice, from foundational test design and standards compliance through to the application of AI-based techniques in safety-critical environments. It addresses DO-178C, ISO 26262, and IEC 62304 as the operational frameworks practitioners work within daily. The central problem is a familiar one for this community: AI tools are entering safety-critical workflows, but those tools were not designed for the verification and compliance demands those workflows require. Maintaining standards' rigour while integrating AI-based approaches is a live engineering challenge across aerospace, automotive, and medical device development.

A companion code repository on GitHub provides working examples. The book is available in e-book and print.

'We wrote the book we wished had existed when we started working at the boundary of AI and safety-critical testing,' said Adam. 'The field is moving fast and the standards community hasn't caught up. We wanted to give practitioners something they could actually use.'

bpbonline.com/products/embedded-software-testing



CERBERUS to evaluate analogue in-memory computing for future systems


Giovanni Ansaloni, EPFL and Sergi Abadal, UPC

In the present post-Moore era, computing systems can no longer be based on ever-smaller, ever-faster and ever-more efficient transistors. Instead, further performance gains, which will be key to coping with the demands of future workloads (especially, but not solely, in the machine learning space) are likely to derive from fundamental shifts in the principles of computer design.

The CERBERUS project, launched in February 2026, addresses this challenge from a multi-faceted perspective. CERBERUS brings together experts in device fabrication, hardware design and integration, system evaluation and machine learning from six institutions across four European countries: Spain (UPC, Barcelona), Sweden (Chalmers, Gothenburg), Germany (AMO and RWTH, Aachen) and Switzerland (IBM, Zurich and EPFL, Lausanne). Together, the CERBERUS consortium studies the applicability of analogue in-memory computing as a disruptive technology for future generations of computing systems.

Our approach will tackle head-on challenges hampering, at present, the mainstream adoption of this disruptive technology. Areas of focus include the study of new, ultra-efficient **memristor devices based on 2D materials**, as well as the development of fast and compact interfaces crossing the boundaries between analogue and digital domains. In addition, we are investigating how thermal effects may influence the accuracy of analogue computation, and novel strategies to minimize thermal noise in co-integrated analogue / digital systems-on-chip, by employing **graphene heat spreaders**. Our full-stack approach enables the abstraction of device-level characterizations, allowing the evaluation of the benefits of CERBERUS technology from a whole-system, whole-application viewpoint.

FURTHER INFORMATION:
cerberus-project.eu
linkedin.com/company/cerberus-project

Principal investigators:
- Sergi Abadal (project coordinator) and Eduard Alarcón (UPC)
- Johan Liu and Per Lundgren (Chalmers)
- Zhenxing Wang and Burkay Uzlu (AMO)
- Renato Negra and Max C. Lemme (RWTH)
- Irem Boybat and Abbas Rahimi (IBM)
- David Atienza, Andras Kis and Giovanni Ansaloni (EPFL)

CERBERUS is funded by the European Union's Horizon Europe research and innovation programme under grant agreement no. 101223271 (EIC Pathfinder Challenges).

Members of the CERBERUS consortium at the project kick-off

ENFORCE: Bringing security orchestration to the computing continuum

As digital services move beyond centralized cloud infrastructures and into edge, far-edge, and internet-of-things (IoT) environments, cybersecurity will need to evolve with them. ENFORCE will address this challenge by developing a standards-driven platform for securing services across the computing continuum. The project will target direct applicability in modern security operations centre (SOC) environments and will combine IoT technologies, advanced 5G connectivity, and edge computing to improve trust, performance, and security in increasingly interconnected infrastructures. Its ambition is not only to detect threats, but also to make complex distributed services easier to deploy, monitor, and protect in real time.

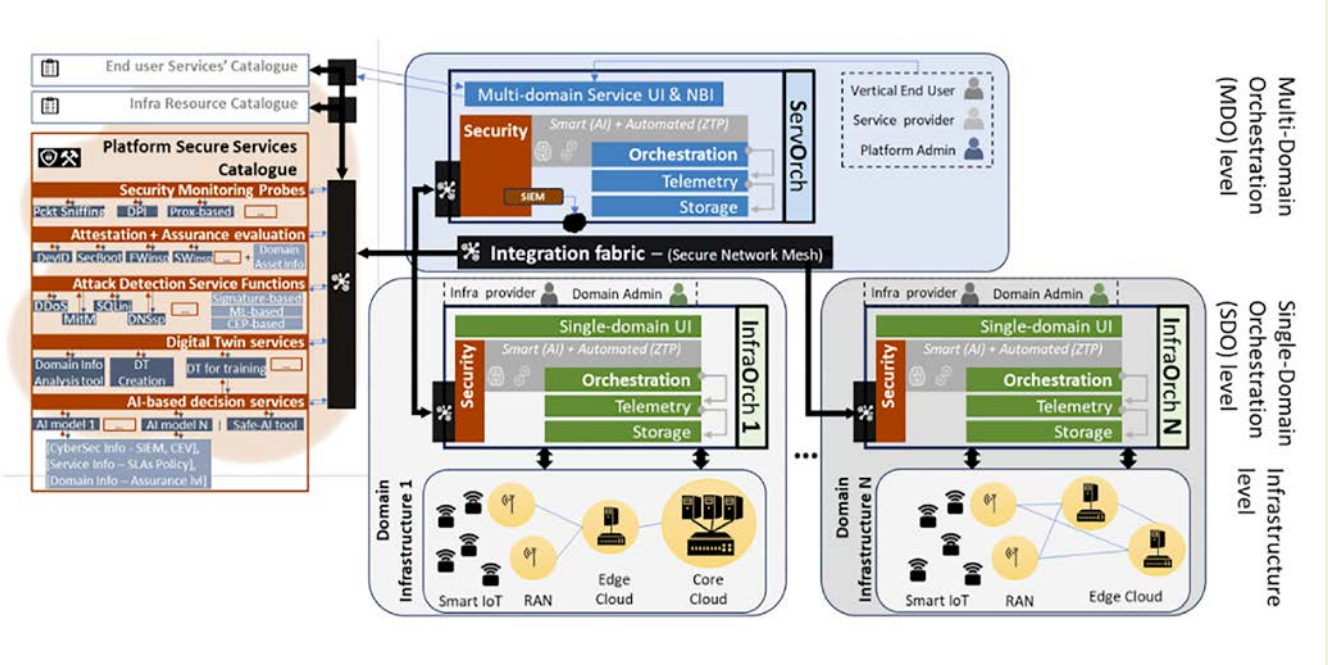
▶

Critical infrastructures are increasingly becoming dependent on distributed services that span heterogeneous domains, from cloud data centres to smart devices at the network edge. Yet these environments remain difficult to secure in a coordinated way. Operators must deal with different trust levels across domains, dynamic infrastructure changes, service-specific service-level agreements (SLAs), software vulnerabilities, and increasingly sophisticated cyberattacks. ENFORCE will tackle these challenges by extending orchestration with security awareness: infrastructure assurance, service-level validation, runtime vulnerability assessment, and automated mitigation will all be treated as first-class functionalities.

ENFORCE takes a holistic approach to security: rather than proposing an isolated security tool, it will integrate security directly into service and resource orchestration. The platform will be designed to assess the assurance level of infrastructure domains as they join the system, validate service deployments against security and performance requirements, and adapt security function chains during runtime. It will combine attestation mechanisms based on trusted platform modules (TPMs) and trusted execution environments (TEEs), monitoring probes, vulnerability scanning, analytics driven by security information and event management (SIEM), and AI-based decision engines. ENFORCE will also explore advanced mitigation mechanisms, including secure TEE migration and automated service reconfiguration, so that security responses can preserve both confidentiality and service availability. The figure below gives an overview of the proposal.

The project is expected to benefit large private organizations, especially telecommunications providers, as well as operators of critical infrastructures and future SOCs. The ENFORCE platform will help them integrate advanced services securely across distributed infrastructure, detect threats faster, coordinate security updates more effectively, and deploy more resilient services. Security professionals and future SOC operators will benefit from cross-domain automation that will reduce the time between detection and response. Beyond this, ENFORCE is also expected to have a broader European impact: its open, standardized approach will support cross-border cybersecurity cooperation, contribute to digital sovereignty, and create exploitable assets and business opportunities.

Validation of this impact is central to the project from the outset. ENFORCE will be demonstrated through three main validation cases: secure domain registration and service provisioning, runtime security updates, and attack detection and mitigation. These will be exercised in two complementary environments: an extended blueprint testing infrastructure provided by multiple academic and non-academic partners, and real telco-site validation environments at Telekom Slovenije in Ljubljana and OTE in Athens. The use cases to be tested will focus on realistic critical-infrastructure settings, including energy systems, smart cities, and logistics. By grounding validation in commercial-grade and operator-grade environments, ENFORCE aims to advance its technologies to technology readiness level (TRL) 7.



The ENFORCE platform

To realize these goals, Members of the ENFORCE consortium brings together 16 partners from eight European Union countries, including telecom operators, small and medium enterprises (SMEs), research institutes, universities, and public cybersecurity authorities. The combination of the partners' deep technical expertise with realistic deployment settings will strengthen Europe's capability to orchestrate secure, trustworthy digital services across distributed infrastructures.

Members of the ENFORCE consortium

EuroTPC: Strengthening Europe's role in global AI for science

The scale and capabilities of artificial intelligence (AI) are rapidly reshaping scientific research across disciplines and placing new demands on the global computing infrastructure that supports it. In addition to compute power, responding to these demands requires coordinated strategy, shared standards and international collaboration. The Trillion Parameter Consortium (TPC) was founded in 2023 to address this challenge. It works to provide an open, community-driven initiative focused on developing large-scale AI models that are trustworthy and built for science.

Initiated by Argonne National Laboratory, RIKEN-CCS, and the Barcelona Supercomputing Center – Centro Nacional de Supercomputación (BSC-CNS), the TPC has rapidly grown to more than 1,300 participants from more than 90 organizations worldwide. Participants collaborate across various thematic working groups on topics ranging from model development and data governance to evaluation frameworks and computational infrastructure. The community-driven nature of the TPC is one of its strengths and is also what makes European engagement both timely and consequential.

EuroTPC was created with this context in mind. Funded under the EuroHPC Joint Undertaking programme, EuroTPC establishes the first coordinated European presence within TPC, connecting Europe's leading high-performance computing (HPC) centres, AI Factories, research institutions and companies to TPC's global activities in a shared direction. The consortium is coordinated by BSC and includes BADW-LRZ (Germany), CINECA (Italy), CSC (Finland), and Neovia Innovation (France). These institutions collectively represent some of the most significant HPC and AI infrastructures in Europe.

EuroTPC works across four main pillars. The first is the establishment of a TPC office in Europe, based at BSC in Barcelona, which will serve as a hub for European engagement in TPC activities and events. The second is the development of a shared roadmap that brings together the agenda of AI Factories, EuroHPC hosting entities, HPC Centres of Excellence and industry into a coherent European direction. For the third pillar, EuroTPC will actively support participation in TPC's scientific activities, from organizing workshops and hackathons to facilitating contributions to working groups. Finally, it will build new international partnerships, deepening existing relationships with TPC partners in the US and Japan, while strengthening connections in Latin America, Asia, Africa and beyond.

For the AI and HPC community, EuroTPC represents a structured way to engage with and shape one of the most ambitious international open AI research initiatives. Europe has the infrastructure, the talent and the institutional experience to play a leading role in shaping AI for science. EuroTPC provides the coordination to act on that potential collectively.

Daisytuner wins third place in T-Challenge finals

In April, Daisytuner – the winner of a HiPEAC Technology Transfer Award in 2024 – reached third place in the T-Challenge finals organized by T-Mobile and Deutsche Telekom in Bonn, out of more than 500 international teams. According to Daisytuner's announcement, Adrian Schmitz and Moritz Timmer from the company presented speedups for object-detection workloads running on graphics processing units (GPUs), achieved without manual tuning. Next, in a live demonstration, they showed how Daisytuner enables performance portability to new hardware, showing the unmodified model running efficiently on a completely different hardware target: a photonic neural processing unit (NPU) from the German company Q.ANT. The model was compiled and executed via Daisytuner's software stack with no changes to the PyTorch code.

Daisytuner's software-optimization platform is designed to automatically compile, tune, and run applications across different hardware processors, without application developers having to adapt their code. The company, which is based in Darmstadt, Germany, built its technology on the foundations of the Data-Centric Programming (DaCe) framework developed in HiPEAC member Torsten Hoefler's Scalable Parallel Computing Lab at ETH Zürich.

FURTHER INFORMATION:
daisytuner.com/news/post-4
bit.ly/T-Challenge_winners_2026

Atakan Aral and Ugur Öztürk awarded third place in University of Vienna Green Inventor of the Year Award

Atakan Aral (HiPEAC member, University of Vienna) and **Ugur Öztürk** (University of Vienna) have been awarded third place in the Green Inventor of the Year Award 2025, announced in February 2026, for their work on a high-performance edge-cloud architecture for climate-resilient railway monitoring.

The proposed system targets landslide and extreme-weather risks affecting railway infrastructure, an increasingly critical challenge for European transport networks. It combines distributed sensing with adaptive edge artificial intelligence (AI) and cloud-based coordination to enable low-latency hazard detection under constrained connectivity and energy conditions.

From a systems perspective, the innovation focuses on efficient data handling across the computing continuum. Data is selectively processed at the edge using lightweight inference pipelines, reducing bandwidth requirements and enabling real-time decision-making close to the physical infrastructure. The cloud layer supports global model updates, long-term analytics, and cross-site correlation, forming a hierarchical architecture for scalable and resilient operation.

A key contribution lies in the co-design of data acquisition, communication, and processing pipelines to minimize data movement while maintaining detection accuracy. This directly addresses bottlenecks in distributed and data-intensive systems, aligning with HiPEAC priorities on data-centric architectures. The modular design further supports deployment across heterogeneous platforms, from low-power internet-of-things (IoT) devices to high-performance cloud backends.

The jury, comprising experts from the University of Vienna's Technology Transfer Office, Austria Wirtschaftsservice Gesellschaft (AWS), the Austrian Patent Office, and the Austrian Research Promotion Agency (FFG), highlighted the system's strong alignment with European Union priorities on sustainable mobility and climate resilience, as well as its clear pathway to real-world deployment.

Building on this recognition, the team is exploring spin-off opportunities to translate the technology into operational solutions for infrastructure operators, with a focus on scalable deployment and integration into existing railway monitoring ecosystems. The official ceremony will take place at the University of Vienna's Green Innovations Event in October 2026.

FURTHER INFORMATION:
transfer.univie.ac.at/univie-inventor-of-the-year-award

David Atienza elected to Academia Europaea

In May, EPFL announced that HiPEAC member David Atienza had been elected to Academia Europaea, the pan-European academy which works to advance excellence in humanities; law; the economic, social, and political sciences; mathematics; medicine; and all branches of natural and technological sciences. David leads the Embedded Systems Laboratory in the School of Engineering at EPFL.

In the announcement, David said that the recognition reflected the work of his team at ESL and the exceptional ecosystem created by EPFL, which provides 'the support and resources needed to allow us to push the limits of computer system engineering and electronic design automation'.

David is the recipient of multiple awards and recognitions, including, most recently, the IEEE CEDA Distinguished Service Award presented at DATE 2026, Design Automation & Test in Europe held in Verona this April.

bit.ly/EPFL_David_Atienza_Academia_Europaea

Sergi Abadal receives Spanish National Research Award

HiPEAC member Sergi Abadal, an associate professor at the Universitat Politècnica de Catalunya–Barcelona Tech (UPC), has received the Matilde Ucelay National Research Award, in the field of engineering and architecture. The award was granted in recognition of Sergi's significant advances in massively parallel computer architectures and in the development of efficient next-generation interconnects. Alongside other honourees, Sergi was recognized in a ceremony in May 2026 presided by the King of Spain and Spanish Minister of Science, Diana Morant.

bit.ly/Premios_Nacionales_Investigación_2025

Photo: © Ministerio de Ciencia, Innovación y Universidades ciencia.gob.es

Onur Mutlu elected AAAS Fellow

In March, ETH Zürich announced that the American Association for the Advancement of Science (AAAS) had named HiPEAC member Onur Mutlu AAAS Fellow for his 'foundational and innovative contributions to computer engineering research, education, and practice, especially in computer architecture and memory and storage systems'.

In the announcement, Onur, who has been a professor of computer science at ETH Zürich since 2015 and leads the SAFARI research group, said: 'This recognition reflects the achievements we have made together with my students and collaborators for more than two decades. It is wonderful to see our work and efforts to have positively and broadly impacted computing systems that are used daily by billions of people in the world and to have positively inspired many researchers and designers to look into new ideas.'

Onur has numerous awards and recognitions and is a fellow of both the IEEE and ACM, as well as a member of the Academia Europaea.

bit.ly/ETH_Onur_Mutlu_AAAS

Hai 'Helen' Li wins IEEE award in circuits and systems

In May, HiPEAC associate member Hai 'Helen' Li (Duke University) was presented the IEEE Circuits and Systems Society Charles A. Desoer Technical Achievement Award, which recognizes individuals for outstanding, consistently generated technical contributions, as well as service to the circuits and systems community. The award was made in recognition of her 'contributions to neuromorphic computing and machine learning acceleration'.

Helen, who gave a keynote talk at HiPEAC 2022, has received multiple recognitions, has authored more than 400 technical publications and holds 79 U.S. patents. She is a fellow of the American Association for the Advancement of Science (AAAS), the Association for Computing Machinery (ACM), IEEE and the National Academy of Inventors.

pratt.duke.edu/news/li-ieee-desoer-award

Yale Patt awarded Osaka Prize

In November, long-term HiPEAC associate member Yale Patt (University of Texas at Austin) was awarded the 2025 Okawa Prize by the Okawa Foundation for Information and Telecommunications, for 'pioneering and outstanding research in high-performance microprocessor architectures, especially involving instruction level parallelism, superscalar processor design, and high accuracy branch prediction'.

The Osaka Prize recognizes people who have made outstanding contributions to the research, technological development and business in the information and telecommunications fields, internationally. Each year one Japanese engineer and one non-Japanese engineer are selected for the award.

Yale's work has been recognized with numerous prestigious awards, including the IEEE/ACM Eckert-Mauchly Award, the 2000 ACM Karl V. Karlstrom Outstanding Educator Award and the Benjamin Franklin Medal for Computer and Cognitive Science from the Franklin Institute.

bit.ly/Yale_Patt_Okawa

Yulia Sandamirskaya awarded International Humanoid Forum Award

HiPEAC member Yulia Sandamirskaya (Zurich University of Applied Sciences–ZHAW) has been recognized with the inaugural International Humanoid Forum Award in the category titled 'Important Academic Contribution'. The award recognizes outstanding achievements and pioneering contributions in the field of humanoid robotics across academia, industry, and applied use cases.

Yulia, who taught at the HiPEAC summer school, ACACES, in 2023, leads the Centre for Cognitive Computing at ZHAW, where her team develops brain-inspired, neuromorphic technology for robot perception, cognition, and control.

humanoid-forum.com/award

Sam Coward wins EDAA Outstanding Dissertation Award at DATE 2026

HiPEAC affiliate member Sam Coward (University College London) was awarded the EDAA Outstanding Dissertation Award in 'Topic 3: New directions in logic, physical design and CAD for analog/mixed-signal,nano-scale and emerging technologies'. Posting about the award, Sam summarized his thesis, titled 'Equality Saturation for Circuit Synthesis and Verification', as follows: 'When you want to make a computer chip design better (faster, cheaper, more power efficient), there is rarely a straight line path. You might need to make the design worse at first, to then uncover really useful optimisations. We got around this problem by efficiently exploring a bunch of design candidates simultaneously using a data structure called an e-graph.'

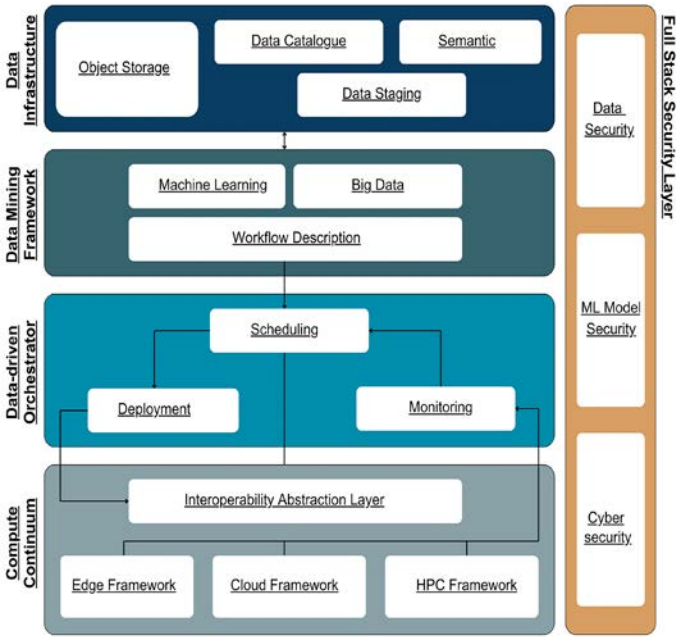
date-conference.com/awards

Read about equality saturation in Sam's article, co-authored with Jianyi Cheng (University of Edinburgh), in *HiPEACinfo* 73
bit.ly/HiPEACinfo73_p22

From scientific instruments to smart-city sensors, the requirements of different data types can vary wildly. In the EXTRACT project, a consortium of 12 international partners led by Eduardo Quiñones, a leading researcher at Barcelona Supercomputing Center – Centro Nacional de Supercomputación (BSC-CNS), set out to investigate computing infrastructure tailored to the needs of different kinds of data. HiPEAC caught up with Eduardo to learn more about the intricacies of leveraging the full compute continuum for data processing, and how EXTRACT's technology was put to the test in two demanding use cases.

'Clearly, you need computing capacity to extract knowledge [from data], but not all data needs the same capacity'

'It's an obvious point, but data on its own means nothing; what is important is the knowledge you extract from the data,' explains Eduardo Quiñones, a leading researcher at BSC who has led multiple research efforts linking the horsepower of high-performance computing (HPC) to the real-time requirements of edge devices. 'Clearly, you need computing capacity to extract this knowledge, but not all data needs the same capacity, nor the same computing resource.' Among the requirements to be taken into account are, for example, interoperability; how data is shared (considering privacy, for example); or whether an outcome is needed in real time. 'One of the key research questions EXTRACT addressed is: how do we design computing infrastructure such that the knowledge you extract from data is generated according to the requirements of different data types in the right computing resource?' says Eduardo.



The EXTRACT architecture spans four main layers of the systems stack

EXTRACT focused not on the data itself but on the compute continuum, investigating what kind of computing resources and abstraction layers were needed to process different kinds of data. 'Most of the time, computing is not considered in its full complexity, but as independent solutions,' says Eduardo. 'For example, most data is processed in the cloud, managed by cloud-specific technologies. There are also edge-specific solutions for processing data at the edge, and solutions tailored to high-performance computing (HPC).' In contrast, the EXTRACT approach aimed to facilitate data manipulation, processing, and analysis across the whole continuum, providing abstraction layers to decide where the data should be computed based on its requirements.

EXTRACT tackled this problem across four main layers of the systems stack:

1. The technologies and features that should be incorporated at **data infrastructure** layer.
2. How the **programmer expresses the workflow or pipeline** that processes the data.
3. What kind of **abstraction** should be used to simplify the experience for the **programmer**, who should be able to see the continuum as **a homogenous environment**, masking the complexity of the **highly heterogeneous computing resources** being leveraged.
4. How the **orchestrator** gets the data requirements and distributes the data to the appropriate place on the computing continuum.

In terms of the data infrastructure, S3 cloud technology was chosen as the common substrate for data storage due to its interoperability capabilities. On top of this object-storage mechanism, two technologies were developed:

- A **caching mechanism** to unify remote and local access, allowing users to access the data in the same way regardless of their location.
- A **mechanism to exploit parallelism**, whereby data is split into chunks depending on its characteristics and distributed accordingly across the available computing resources. 'This is fundamental when dealing with massive amounts of data which need to be processed on cloud or HPC resources,' explains Eduardo. Such extreme data requirements are exemplified by the project's radio-astronomy use case: data received from hundreds of antennas is filtered in real time and significantly compressed (by a factor of 100) before being stored on S3. 'Our distribution mechanism takes this data and, based on its characteristics, divides it into chunks to be processed in parallel,' says Eduardo. 'This results in a science product upon which astronomers can conduct their research.'

On top of this data infrastructure, EXTRACT used different programming technologies to express the data-processing workflow. The serverless technology Lithops was chosen as, in common with other 'function as a service' software, it allows users to focus exclusively on their desired function, which in this case was ensuring that the data is in the right place, Eduardo notes. 'Lithops wakes up quickly, executes the function, and terminates, which makes it particularly suitable for parallel execution.'

Python was the programming language used, while EXTRACT also incorporate low-level languages and programming models such as OpenMP, CUDA, and MPI for HPC execution, as well as experimenting with BSC's in-house COMPSs programming model.

The cloud container-management software Kubernetes played a key role in abstracting the computing resources to enhance programmability. 'One of the key innovations in EXTRACT was to match the execution model of workflows to a Kubernetes "multicluster" of edge, cloud and HPC resources, as outlined in a recent paper accepted for the IEEE World Congress on Services this year, so that the function is automatically mapped to the relevant computing resources,' says Eduardo. 'The multicluster abstraction layer combined with the workflow execution model allows computation to be sent to the specific resource so that when the data arrives it is ready to execute the function, managing dependencies between functions and network complexity.'

Members of the EXTRACT project team at HiPEAC 2025

Another key component developed by the EXTRACT project is the orchestrator: this considers functions and decides where to deploy and where to execute the computation along the edge-cloud-HPC continuum, based on application requirements and taking into account the availability of different computing resources. Responding to the major advancements in artificial intelligence (AI) since the project proposal was developed, EXTRACT researchers also investigated how agentic technologies could be used to configure compute resources for efficient execution. 'The agentic concept implies that each agent includes a language model which reasons about the best configuration each resource should include to execute the workflow,' explains Eduardo. 'As well as testing how to scale resources within a cluster using the cognitive orchestrator, we've also been experimenting with agents to configure the communication parameters to stream data between clusters. Deploying this kind of technology fits perfectly into EXTRACT. All the complexity under the hood is hidden by this distributed-infrastructure monitoring layer, using agentic AI not only at the application level, but also to manage resources.'

As well as being used to serve data, the computing continuum is also exploited for AI models, says Eduardo: 'AI models evolve in a constant training loop. Trained on large machines – such as HPC facilities – once they reach a suitable performance level, they can be moved to the edge of the continuum for deployment – which means that you need model-serving functions.'

Extreme data needs

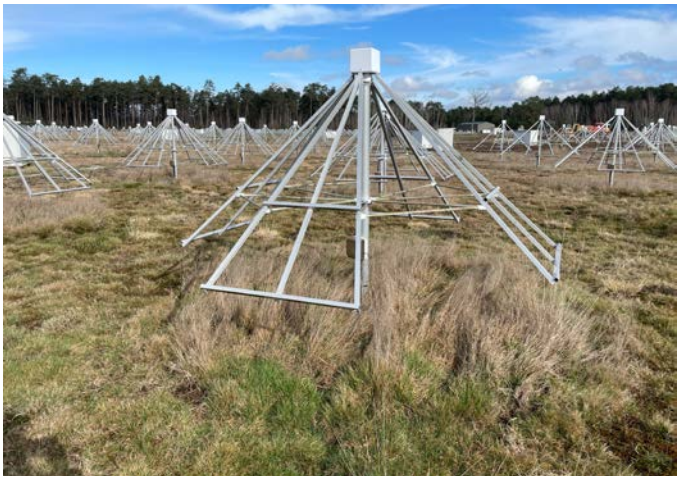
EXTRACT put these technologies to the test in two demanding use cases. For the first, the radio-astronomy use case, EXTRACT technology was used to develop data-mining workflows that effectively reduce the huge amount of raw data produced by NenuFAR radio telescopes by a factor of 100. This allowed high-quality datasets to be populated and made available to

the astronomy community (through the European Open Science Cloud–EOSC).

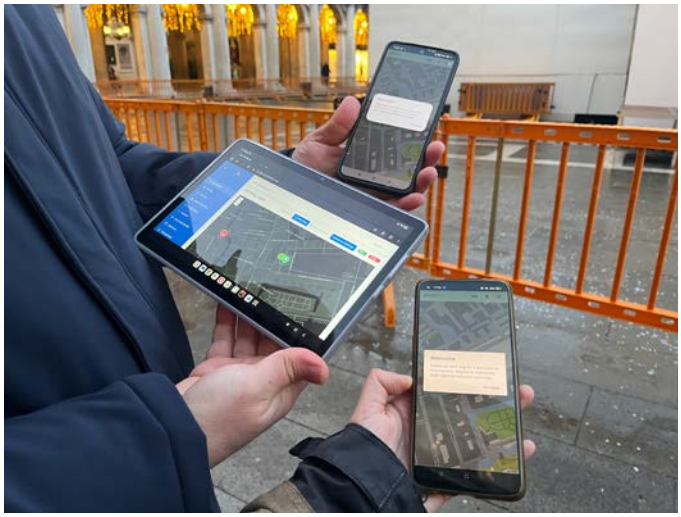
'Antennas from the Square Kilometre Array Pathfinder are constantly receiving signals from the universe, but with the technology we have today, we can only process around eight hours of data. For this use case, we needed to optimize the edge devices to compress the data with AI models using low-level programming languages like OpenMP or CUDA. Real-time data is then stored in S3 files, before being moved to the cloud where it can be processed in the parallel capabilities of the EXTRACT technology. We also addressed the need to constantly improve the AI models with model serving solutions.'

The second use case used the EXTRACT platform to develop, deploy and execute a data-mining workflow to generate personalized evacuation routes (PERs) in the event of an emergency, which would be delivered directly to the individual displayed in a mobile app. To do this, the EXTRACT platform had to process and analyse data composed of Copernicus and Galileo satellite data, data from internet of things (IoT) sensors installed across the city, 5G mobile signals, all of which flowed into an urban digital twin.

'While not strictly speaking an extreme-data use case, in the sense that the amount of data is not massive, this use case exercises an important part of EXTRACT: real-time operation,' explains Eduardo. 'The AI model needs to be trained on HPC facilities with reinforcement-learning methods, based on multiple simulations generated by the urban digital twin in the cloud. It receives real-time data from multiple sensors: in the case of the EXTRACT test city, Venice, these include sensors for flooding, fire, and geolocation data from individuals' mobile phones.' Based on this data, EXTRACT has developed a semantic data information ontology, allowing the AI model to reason about the data.



NenuFAR radio telescopes in Nançay, France, generate huge amounts of raw data



The PER use case incorporated data from multiple sensors to help users navigate Venice's dense maze of streets

The digital twin provides a snapshot of the current status of the city, which is then used to generate different simulations. Although requiring HPC facilities for the training phase, when an emergency occurs the trained model is then deployed at the edge, next to the antennas – and is thus a good fit for the 5G paradigm, says Eduardo.

However, he notes that finding a model capable of determining PERs is very complex. 'You need to create a representation of the city that includes sufficient information for the AI agents to be able to route people to a safe spot. While Venice is not a large city, it has a dense concentration of streets.'

Apart from the intense requirements of the model, the PER use case also posed significant challenges in terms of cybersecurity; indeed, the distributed nature of computing resources used in EXTRACT as a whole meant that security was a major concern. 'Although we didn't plan to advance beyond the state of the art in security, we were conscious that it is a fundamental requirement for such a personalized use case,' says Eduardo. 'We deployed mechanisms to ensure data privacy, and also analysed the code to ensure that there were no leaks in the deployments on different computing resources.'

The technology developed by EXTRACT could also be used for other used cases, Eduardo adds. 'Any use case that you can describe with a workflow, and where data is gathered from distributed data sources and processes across some form of compute continuum, can be executed using EXTRACT,' he says. 'Examples include management of the power grid, logisitics, and smart agricuture.'

FURTHER INFORMATION:
EXTRACT project website extract-project.eu

Supporting data needs from the edge to HPC

How LEXIS unlocks advanced computing while SMARTEDGE puts the privacy into edge data

As part of our special feature on data, we find out how LEXIS allows workflow orchestration and distributed data management on advanced computing facilities, while SMARTEDGE provides a solution to the thorny problem of deriving insight from crowd-sensed data while maintaining anonymity.

LEXIS Platform: A gateway to supercomputing, quantum and AI technologies

Martin Golasowski, IT4Innovations

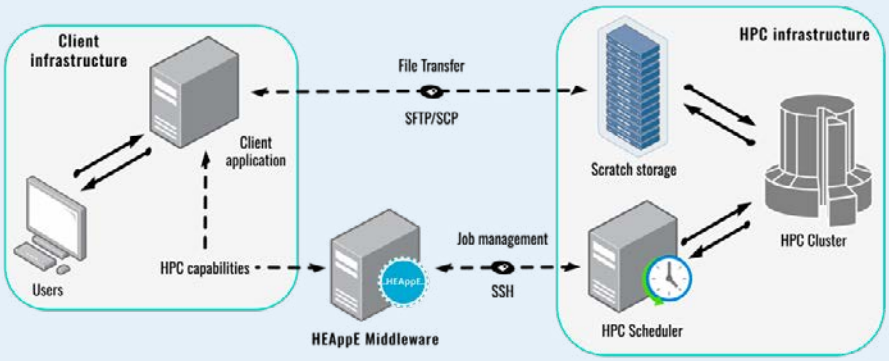
For decades, entering the world of high-performance computing (HPC) was largely limited to a small group of experts fluent in command-line interfaces. With the LEXIS Platform and its integration into the EuroHPC Federation Platform, this situation is changing. The ambition is to offer scientists, startups, and industrial users supercomputing power that can be used directly from a web browser with a single click.

Developed since 2019, the LEXIS Platform represents a major shift towards easy and secure access to advanced computing resources. Instead of relying on complex Secure Shell logins and manual data transfer, users can manage the entire lifecycle of a computation through an intuitive web interface. They simply upload their data, define their workload as either a custom job script or a container and the platform takes care of the rest – from transferring data and launching the jobs across multiple clusters to real time monitoring and delivering the results. Apart from a generic web-based interface, the platform also provides several domain-specific graphical interfaces, such as an AI-oriented interface tailored for efficient artificial intelligence / machine learning (AI / ML) model lifecycle management, including training and inference services.

The platform's capabilities are enabled by several key technologies developed at IT4Innovations National Supercomputing Center in the Czech Republic:

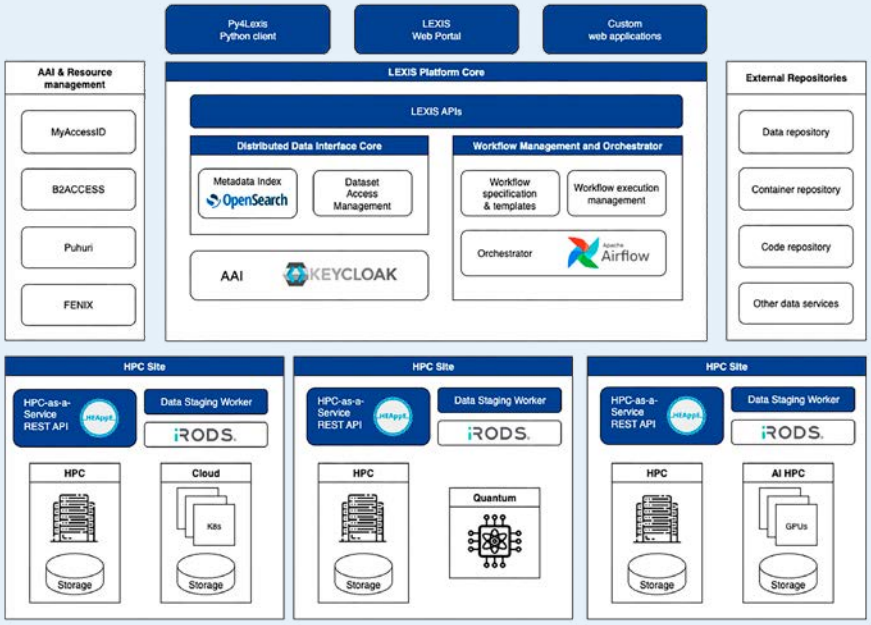
- **HEAppE (HPC as a service)**
  This high-end application-execution middleware forms the bridge between modern web interfaces and traditional HPC environments. It exposes a REST application programming interface (API) that handles user authentication through OAuth2 standard, offers various identity mapping strategies, providing a common mechanism for job submission across locations.
- **Distributed Data Interface (DDI)**
  This software solution provides crucial data-management and transfer capabilities for the platform, offering managed **iRODS** zones deployed on local storage, a centralized metadata index based on OpenSearch and asynchronous data transfers between data repositories and computing resources.
- **Workflow orchestrator** and builder based on Apache Airflow
  The workflows are orchestrated by an Apache Airflow with two LEXIS Platform-specific extensions. The LEXIS provider library contains the building blocks, individual tasks calling **HEAppE** and **DDI** APIs, which are then used to form workflow patterns and user-specific workflows using easy and declarative YAML-based description (LEXIS Workflow Definition). This capability is provided by a LEXIS Platform specific Airflow plugin. A web-based workflow builder for drag-and-drop definition of the workflow patterns is also available.

Thanks to its intuitive interface and powerful data-management capabilities, the LEXIS Platform has been used to power several ambitious European projects. ▶



LEXIS simplifies access to advanced computing resources

- In the LIGATE project, the LEXIS Platform supported large-scale drug-discovery workflows. The pharmaceutical company Dompé uses the proprietary LiGen software, co-developed with Politecnico di Milano and CINECA. Because intellectual property constraints prevented the sharing of LiGen binaries or source code, LEXIS Platform provided a secure way to expose the workflows to external users without revealing the underlying software.
- LEXIS Platform is used as a cornerstone for data and computing task management in several application cases of the EXA4MIND project, including the systematic improvement of molecular-simulation accuracy, large-scale automated annotation and evaluation of automotive camera recordings, and smart farming / viticulture scenarios that combine sensor data with satellite imagery.
- The OpenWebSearch.eu initiative built an open and transparent European infrastructure for web crawling, indexing, and data sharing. The project uses European computing resources for large-scale crawling, data management, and processing through the LEXIS Platform.



- The BioDT project, which supported the development of digital twins for biodiversity research, also benefited from the LEXIS Platform. The platform allowed execution of workflows containing HPC tasks from a custom user interface and it was used to execute computation of new data products.

The strategic relevance of the LEXIS Platform is evident in its integration into the recently released EuroHPC Federation Platform (EFP). This major initiative seeks to unify Europe's fragmented supercomputing landscape under a single, user friendly service with common authentication and authorization infrastructure (AAI) and easy-to-use graphical user interfaces and advanced tooling for power users. Key components originating from the LEXIS Platform will form the backbone of workflow orchestration and distributed data management. Through a common service, users will gain unified access to a diverse ecosystem of systems – from classical supercomputers to emerging quantum technologies and AI factories.

lexis.tech

Anonymous reputation update as an edge service

Alex Shafarenko, University of Hertfordshire

smartedge

One of the most challenging problems in crowdsensing stems from the fact that smart sensors may be associated with human actors, and that sensor reports may inadvertently capture protected personal data as an unintended byproduct of submitting a situation report. Even if the human is protected by a pseudonym and their real identity is not exposed, correlating reports submitted under the same pseudonym can effectively de-anonymize the actor.

Consider, for example, a city-wide traffic-monitoring system in which road users anonymously submit traffic data: distance to the vehicles in front and behind, current speed, and precise location. Such data are valuable for traffic control, enabling smart traffic lights and variable speed limits for predictive congestion management. The byproduct, however, is that edge systems can store each pseudonym's trajectory over time.

Visiting specific locations – offices, schools, gyms, doctors' surgeries, and so on – at regular times progressively increases the certainty with which an actor can be identified, and may quickly

prove sufficient to break their anonymity entirely. There cannot be many people who drive their children to school X, visit gym Y at a particular time, and patronise the same set of shops. Even if deliberate surveillance is unlikely once anonymity has been broken on a single occasion, a potentially unlimited amount of personal trajectory data becomes available to the curators of the traffic control system.

Since crowdsensing critically depends on the willingness of human actors to participate, it is in data managers' best interests to convince the public that providing sensor data does not amount to unwittingly spying on oneself. It is possible, of course, to rely on trust between the sensor and the control system – whether for traffic, environmental, health, or other forms of crowdsensing – to protect privacy. However, as the Finnish authorities' attempt to introduce a rolling per-mile road charge by installing GPS locator devices on cars demonstrates, the general public is deeply apprehensive of such intrusions and tends not to trust those who curate their data.

A further problem is that fully anonymous users may include bad actors attempting to manipulate the monitoring system by submitting false reports for personal gain. A vehicle, for instance, could falsely signal a traffic jam on the approach to a set of lights in order to keep them green in its direction of travel at the expense of cross traffic. The standard solution is to weight reports from different actors according to their reputation. Actors whose reports prove accurate after the fact have their reputation increased, while those who abuse the system see it fall rapidly to zero.

To build a reputation, reports from the same actor must be linkable – yet to protect the actor's identity, such linkage must be impossible. This appears to be an insuperable problem.

A solution exists when it is not the reports themselves, but time-stamped reputation values assigned to the actor, that are accumulated – while the reports are assessed strictly on an individual basis within the context of the sender's verified current reputation. A registration authority (RA) holds reputation data under a pseudonym, but not the reports themselves. An actor obtains a reputation certificate from the RA and submits it to a separate reputation server (RS) alongside the sensor report. The RS collects reputation-marked reports from multiple users, infers the situation from them, and returns a reputation update coupon to each sender. If the certificate and coupon are provably free of any identity information, the actor can present the coupon to the RA and receive an updated certificate – without the RA gaining any knowledge of the report, even if it were secretly colluding with the RS.

This is the core idea of the zero-trust protocol for anonymous reputation update. Zero-trust imposes strict constraints on the servers: they cannot, for example, use random data in messages, since randomness cannot be proven – random values could secretly carry identity information, compromising unlinkability.

The protocol must clear two technical hurdles. First, the certificate and coupon must be unique (unclonable): the certificate must not identify the actor, and the coupon must be linked to the actor without the RS being aware of that link. This is achieved through a technique known as blind signature with public metadata. Second, the RS requires an authenticated channel for report submission, which in turn requires establishing a shared secret

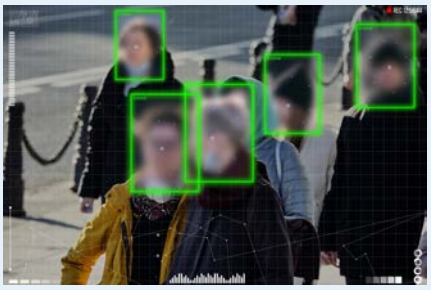


Photo: gesrey | stock.adobe.com

at the session level. Both techniques are computationally demanding for an internet-of-things (IoT) smart sensor running on autonomous power.

This problem was addressed in the context of two use cases within the European Union project SMARTEDGE, supported by two industrial partners: EMC (Slovakia), contributing a healthcare crowdsensing scenario, and Conveq (Finland), contributing a traffic control scenario – with myself acting as a consultant. We have developed a new protocol based on a novel approach to RSA blind signatures with public metadata, and on the Guy Fawkes protocol technique. Our case studies show that it is possible to construct a sealed anonymizer – in the form of a key fob, powered by a non-rechargeable battery – that operates for at least three years as an intermediary between car sensors and the telecommunications infrastructure for data transfers between the vehicle and the RA/RS.

FURTHER READING:

A. Shafarenko, 'A zero-trust swarm security architecture and protocols', Cryptology ePrint Archive, Paper 2025/1105, 2026-03-31
eprint.iacr.org/2025/1105

This work was supported by the European Union's Horizon Europe research and innovation programme under grant agreement no.101092908. Views expressed are those of the author(s) only and do not necessarily reflect those of the European Union.

Manufacturing and healthcare settings across Europe are potent sources of valuable data, yet accessing this data and acting upon the insights it provides in real time is significantly complicated by protocol fragmentation and device silos. In this article, Sravani Thota, and Deepak V. Katkoria, of the Austria-based company Logiicdev, introduces KiiGen, a research-authenticated platform based on field-programmable gate arrays (FPGAs) which harmonizes data and uses artificial intelligence (AI) to solve critical industrial problems in real time.

How KiiGen™ turns industrial data chaos into deterministic intelligence

In today's manufacturing context, equipment such as factory robots, industrial programmable logic controllers (PLCs), sensors and legacy machines constantly generate rich operational data. However, this data is trapped behind protocol silos, non-deterministic networks, and insecure, ad hoc integrations. Walk into most factories today and you'll find robots communicating with EtherCAT, conveyor systems working with Modbus, PLCs communicating via PROFINET and radio-frequency identification (RFID) gateways using the MQTT messaging protocol. This is inefficient, potentially unsafe, and deeply at odds with the promise of Industry 4.0.

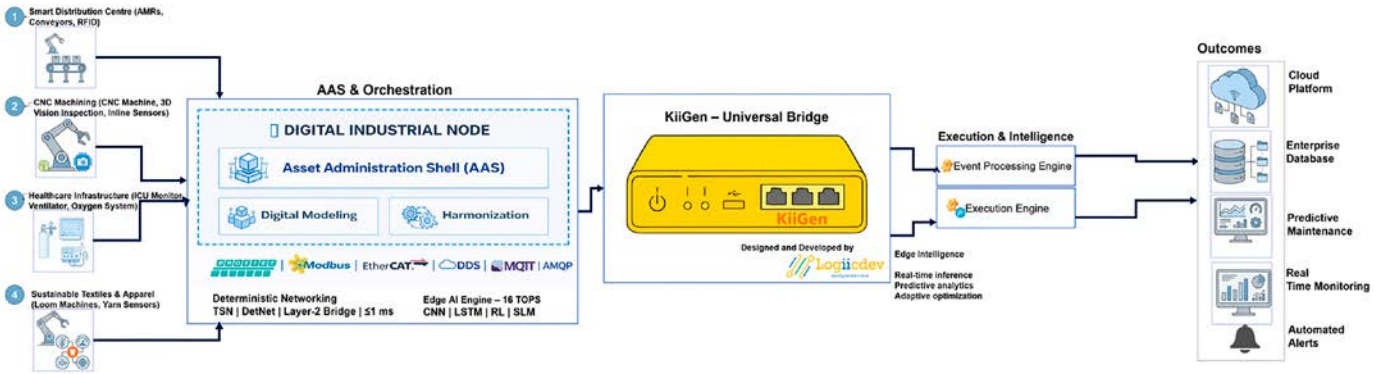
Developed by Logiicdev in Graz, Austria, KiiGen is a modular edge platform engineered to resolve this problem at its root. Rather than requiring factories to replace legacy equipment, KiiGen acts as a universal bridge layer, ingesting heterogeneous industrial data streams, harmonizing them in real time, and delivering deterministic, AI-enriched outputs to dashboards, PLCs, enterprise databases and cloud platforms.

The data problem at the edge

Factories produce huge volumes of data, such as sensor readings, machine logs, and process telemetry. The challenge for industrial-data management is to overcome issues of latency, fragmentation, and actionability. Timing is paramount: data generated by a computer numerical control (CNC) spindle must influence a cutting parameter adjustment in a matter of microseconds, for example. If protocols are incompatible and systems are siloed, it is impossible to achieve this kind of real-time result.

KiiGen addresses this at the hardware-software intersection. Its FPGA-based protocol bridge translates between PROFINET, EtherCAT, Modbus, CAN, PROFIBUS, MQTT, and DDS simultaneously, enforcing time-sensitive networking (TSN) and layer-2 bridging to guarantee end-to-end latency of 1 millisecond or less. By implementing deterministic packet scheduling in reconfigurable hardware – as opposed to a software polling loop – KiiGen ensures time-critical industrial control, essential for closed-loop CNC feedback, robotic coordination, and life-critical medical-device orchestration.

More than just a gateway, KiiGen acts as an orchestration node that understands the timing, priority and semantics of traffic flows, enabling scheduled traffic, quality of service (QoS) prioritization and the safe coexistence of control, monitoring and AI data streams on the same physical network. As a result, data can be moved, fused and redistributed in real time between



KiiGen's FPGA-based bridge translates between protocols simultaneously

PLCs, edge AI engines, manufacturing execution systems (MES) systems and cloud platforms without violating deterministic constraints.

From protocol bridge to edge AI brain

Once the data is harmonized, KiiGen uses AI to extract insights from the data or integrate AI-generated action to data. Its edge AI engine – performing 16 tera operations per second (TOPS) – is embedded in the orchestration layer for machine-level inference which takes place entirely at the edge, without the need to send data to the cloud. Various kinds of machine learning are used, including:

- convolutional neural networks (CNNs) for visual inspection
- long short-term memory models, a kind of recurrent neural network, for time-series anomaly detection
- reinforcement agents for throughput optimization
- predictive-maintenance models

Ensuring that the AI takes place exclusively at the edge preserves data sovereignty, eliminates cloud latency, and allows factories to operate autonomously even when connectivity is degraded. This is crucial for compliance in sectors handling, for example, sensitive production intellectual property (IP), healthcare data, or regulated processes. The platform also integrates Asset Administration Shell (AAS) models to create structured digital-twin representations of each connected asset, enabling legacy machines to be monitored and optimized identically to brand-new equipment, without modification.

Impact across sectors

KiiGen's architecture is domain-agnostic by design. In CNC machining, real-time 3D vision inspection at sub-millisecond cycle times enables zero-defect production. In sustainable textile manufacturing, computer vision achieves 99% yarn-break detection accuracy and reinforcement learning drives throughput gains of up to 40%, while sensor-driven analytics automate environmental and social and governance (ESG) and carbon-border adjustment mechanism (CBAM) compliance reporting. In smart logistics, multi-protocol synchronization of autonomous mobile robots (AMRs), conveyors, and RFID readers eliminates vendor fragmentation while enabling safe human-robot collaboration.

Perhaps the most compelling application is healthcare. KiiGen's secure orchestration layer integrates intensive-care unit (ICU) monitors, ventilators, and oxygen delivery systems using DDS and MQTT harmonization, with deterministic networking ensuring that latency spikes – potentially fatal in life-critical environments – are architecturally impossible, and the system maintains 24/7 stability without external cloud dependency.

Security as a first-class data concern

Industrial data security has historically been an afterthought bolted onto systems designed without network connectivity in mind. In contrast, KiiGen integrates trusted platform module (TPM) secure boot, authenticated firmware updates, and encrypted communication channels as architectural primitives, not add-ons. All operational datasets are anonymized and processed under international industrial data-governance standards, ensuring the data pipeline KiiGen creates is not itself a new attack surface.

Conclusion

The data bottlenecks plaguing industrial environments are an architectural problem rooted in decades of protocol fragmentation and siloed thinking. KiiGen's contribution is a coherent edge-layer answer: deterministic, secure, AI-capable, and deployable onto existing infrastructure. As manufacturing moves toward Industry 5.0's vision of adaptive, human-centric, sustainable operations, the ability to unify industrial data in real time will be a defining engineering challenge of the decade. KiiGen ultimately positions data, not devices, at the centre of industrial innovation. By turning heterogeneous, insecure and non-deterministic manufacturing environments into a unified, secure and deterministic data infrastructure with embedded AI, it directly addresses key HiPEAC themes in data movement, processing, cybersecurity and privacy, while opening a practical path from Industry 4.0 to Industry 5.0.

KiiGen originated in the Ros6BUSBridge research project, a collaboration between CEA-List (France) and Logiicdev, which pioneered deterministic ROS2-to-industrial-bus interoperability and multi-protocol translation frameworks. The underlying research was peer-reviewed and published at ETFA 2025, the IEEE 30th International Conference on Emerging Technologies and Factory Automation (see 'Further reading', below). The technology stack combines open-source frameworks including TensorFlow Lite, PyTorch Mobile, and Docker with Logiicdev's proprietary real-time protocol translation engine and custom HDL deterministic packet logic.

FURTHER READING:

Q. -D. Nguyen, D. Rammal, C. Gaston, D. V. Katkoria, A. Lapitre and S. Dhouib, 'Towards Bridging Industrial Ethernet Networks: Protocol Translation and Runtime Verification', 2025 IEEE 30th International Conference on Emerging Technologies and Factory Automation (ETFA), Porto, Portugal, 2025, pp. 1-4, doi: 10.1109/ETFA65518.2025.11205755

ieeexplore.ieee.org/document/11205755

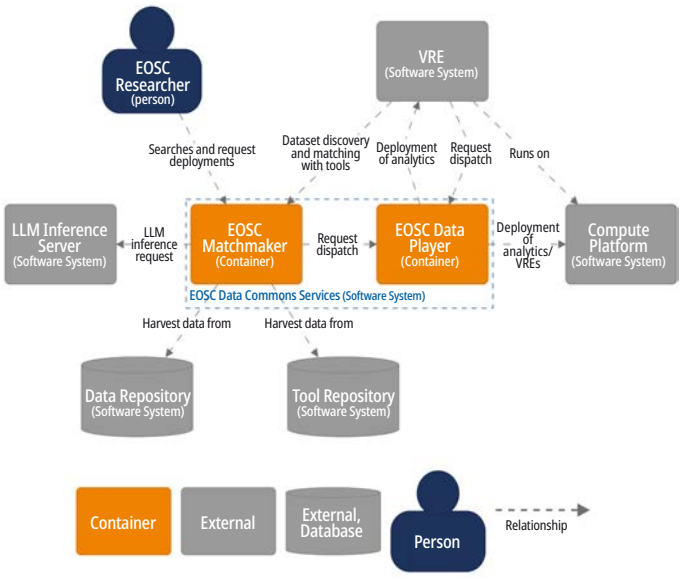
Europe benefits from world-class research infrastructure, yet the distributed nature of scientific resources can impede data sharing and reuse. Funded by the European Union, EOSC Data Commons is working on a federated compute and data continuum to turbocharge scientific discovery. In this article, Xavier Salazar (EGI Foundation) introduces the project, its objectives and its main services.

EOSC Data Commons

Building a federated data infrastructure for European research

Research across Europe is becoming increasingly data intensive. From large scientific instruments and simulations to digital humanities and social science datasets, researchers now work with volumes of data that are growing in scale, complexity, and diversity. The challenge is not only collecting this data but also making it usable: data is often distributed across repositories, stored in different formats, described with inconsistent metadata, and governed by incompatible access policies. These barriers hinder discoverability, reuse, and cross-disciplinary analysis.

EOSC Data Commons aims to reduce this fragmentation and thereby prepare the next generation of research data infrastructure for Europe. Launched in April 2025, the project contributes to the European Open Science Cloud (EOSC) by connecting data repositories, virtual research environments, and common analytical tools and services into a federated compute and data continuum. Rather than centralizing all data in one place, the goal is to make distributed resources work together more effectively while preserving the autonomy of individual providers.



EOSC Data Commons high-level context architecture

EOSC Data Commons allows researchers to discover relevant data, access the necessary tools, and run analyses without wasting time dealing with incompatible interfaces, missing metadata, or manual data transfer. It brings data and computation closer together and makes research infrastructure more interoperable across domains and borders.

EOSC Data Commons offers two complementary services:

- **EOSC Matchmaker** is a discovery layer for datasets, tools, and services. It helps researchers identify relevant resources across scientific domains and connect them with analytics services that can be deployed near the data. This is especially important when dealing with large or distributed datasets, where moving the data is slower, more expensive, and sometimes simply impractical. By deploying modular, technology-specific crawlers, it systematically harvests and harmonizes heterogeneous metadata from thematic, national, and institutional repositories (supporting standard protocols like OAI-PMH and DataCite out of the box).

  The Matchmaker moves beyond keyword searches by implementing large language models (LLMs) via the Model Context Protocol (MCP) to translate natural-language user queries into structural SPARQL queries over a virtual knowledge graph. It assesses dataset FAIRness (i.e. findability, accessibility, interoperability, and reuse) and automatically pairs discoverable data assets with compatible analytical tools. The definitive output of this service is a machine-actionable package for processing datasets, an interoperable standard combining file-level metadata wrapping (using BagIt specification) within a comprehensive RO-Crate package.

- **EOSC Data Player** provides the interoperability layer that allows these services to work across multiple infrastructures. It focuses on harmonized application programming interfaces (APIs), shared metadata specifications, and common mechanisms for authentication, authorization, and provenance tracking. In practice, data and tools from different providers can be combined more easily while keeping track of where the data came from, how it was processed, and under what conditions it was accessed.

  Acting as the execution core, the Data Player ingests the generated RO-Crate processing packages to orchestrate computation across the European compute continuum – from interactive environments to high-performance computing (HPC), high-throughput computing (HTC), and cloud platforms. A central dispatcher component validates incoming execution requests and coordinates tasks between pluggable data access layers and compute engines.

  To eliminate vendor lock-in and keep applications cloud agnostic, the platform uses open-source deployment tools like the Infrastructure Manager (IM), which automates the provisioning, installation, and on-demand contextualization of complex virtual infrastructures on underlying infrastructure-as-a-service (IaaS) clouds via TOSCA templates. Workloads are then submitted through an extensible plugin framework supporting mature runtime environments – such as Galaxy, BinderHub, JupyterHub, ScienceMesh, and OSCAR for highly parallel, event-driven serverless data processing.

Key challenges

EOSC Data Commons addresses several key issues relating to pan-European data infrastructure. First, it takes into account the fragmentation of European research infrastructure, characterized by different repositories and service providers across countries and domains operating with their own governance, standards, and technical environments. To solve this, EOSC Data Commons builds on a distributed architecture and standardized interfaces and tools to connect them. This federated approach allows each provider to retain control over its own resources while operating analytic tools within a shared ecosystem.

Second, the project is working to ensure that proper metadata is provided, since without rich, consistent descriptions, data cannot be reliably discovered, understood, or reused. EOSC Data Commons makes extensive use of semantic technologies to improve discovery and contextualization and to help link datasets, software, workflows, and other research artefacts, so that machines can interpret them and services can reuse them.

Third, the project adopts community standards such as RO-Crate, which provides a machine-actionable way to package research artefacts together with their metadata. Standards like this are important because they reduce ambiguity and make it easier for tools to automate tasks such as validation, packaging, exchange, and preservation. This allows researchers to spend less time on manual curation and more time on analysis.

Finally, as datasets become larger and more complex, manual exploration alone is no longer enough. The EOSC Matchmaker uses artificial intelligence (AI) to simplify search and analysis at scale, helping researchers identify patterns, navigate large collections, and extract useful information more efficiently.

Real-world implementation

EOSC Data Commons services are being developed and validated through scenarios from life sciences, social sciences, environmental research, physics, and other domains. By basing the development of the services in real research workflows, where data formats, access rights, and analysis requirements vary across fields, the project is addressing real-world interoperability problems. Services are already available for testing with three use cases (DABAR, DANS, and SWISSUbase) providing data to EOSC Matchmaker, which leverages a set of deployment tools, orchestrators, virtual research environments (VREs), cloud and container management frameworks, and data access services operating across the compute continuum.

To further this work, EOSC Data Commons has an open call for external organizations to become early adopters of the services. External parties are welcome to participate in the testing and co-design of EDC services and contribute feedback to their release, strengthening the broader EOSC landscape in the process.

A shift for European science

The project's wider ambition is to support the full research-data lifecycle, including data deposition, discovery, access, analysis, and long-term preservation. Improving interoperability and reducing data movement, EOSC Data Commons will help make research outputs more FAIR. It will make distributed data easier to find, combine, and reuse, helping create the conditions for more collaborative, more reproducible, and more data-driven research across Europe. This has major implications for European science: beyond simply improving efficiency, it changes what kinds of research become possible.

FURTHER INFORMATION:
EOSC Data Commons eosc-data-commons.eu

In 2020, the European Union set out its strategy for data, with the aim of harnessing the value of data for Europe. One key area where data sharing can make a significant impact is government transparency. In this article, CEDAR project Vice Coordinator Theodoros Semertzidis, (Centre for Research and Technology Hellas–CERTH), and Dissemination Manager Francesco Osimanti (The Lisbon Council) explain how CEDAR is using advanced computing technologies to transform diverse data sources into a cross-border intelligence layer that ensures accountability and effectiveness in public spending.

Transforming EU governance

The CEDAR project and the dawn of a data-driven public procurement era

Public procurement is the engine room of the European economy, accounting for approximately 14% of the European Union's gross domestic product (GDP) – over €2 trillion annually. Yet for decades this vast expenditure has been managed through fragmented, localized systems that often lack transparency, making them vulnerable to inefficiency, fraud, and corruption. Enter the **CEDAR project** (Common European Data Spaces and Robust AI for Transparent Public Governance), a Horizon Europe initiative (2024–2026) designed to revolutionize the full lifecycle of procurement data management.

By leveraging advanced artificial intelligence (AI) and the transformative power of Common European Data Spaces (CEDS), CEDAR is going beyond digitizing records to build a cross-border intelligence layer that ensures that public money is spent accountably and effectively.

From fragmented data to FAIR intelligence

The primary hurdle in public governance has never been a lack of data, but rather its 'siloed' nature. Procurement notices, contract awards, and invoices are often scattered across thousands of regional portals in varying formats, from structured databases to unsearchable PDFs.

"The primary hurdle in public governance has never been a lack of data, but rather its siloed nature"

CEDAR addresses this through a robust **data management cycle** rooted in the FAIR principles (findable, accessible, interoperable, and reusable). The project implements a sophisticated **data operations (dataOps) and machine learning operations (MLOps) pipeline** that automates the ingestion, cleaning, and harmonization of data from diverse sources (text and media). Using a focused ontology and a knowledge graph that links together tenders, bids, and actors, CEDAR creates a 'common language' for procurement, ensuring that a contract award in Italy can be semantically compared to one in Slovenia or Ukraine.

Since different organizations are at different levels of maturity in terms of digitization, this cycle doesn't start with digitized records and it doesn't end with data storage. CEDAR provides tools and services for digitizing scanned documents using AI, processing and cleaning the data to be fitted to the knowledge graph, and data stored in appropriate pseudo-anonymized repositories. Finally, through secure connectors and application programming interfaces (APIs), the project ensures that high-quality, pseudonymized datasets are ready for the next crucial stage, i.e. deep analysis via AI to detect misuse indicators and further investigation alerts. CEDAR here is at the heart of the CEDS mission to produce and analyse data that improve intelligence analysis.

Advanced AI for continuous public integrity checks

At the heart of the CEDAR architecture lies a suite of advanced AI models designed to act as a sentinel. Unlike traditional oversight, which often relies on manual audits of a tiny fraction of contracts, CEDAR's AI analyses 100% of the available data in real time.

The project utilizes **explainable AI (XAI)** and graph-based data modelling to detect patterns that are invisible to the human eye. The system operates on two levels of triggers:

1. **Simple indicators**
   These are the 'red flags' often associated with compliance. They include indicators like 'single-bidding' (where only one company applies for a tender), 'purchase splitting' (breaking a large contract into smaller ones to avoid oversight thresholds), or 'negotiated procedures' without prior publication.
2. **Complex indicators**
   These are where CEDAR's AI truly shines. By integrating external data sources – such as company registries and historical bidding behaviour – the system identifies hidden 'collusive networks'. It can detect **bid rotation** (where companies take turns winning), **price manipulation**, and **tailored tenders** (where requirements are so specific they can only be met by one favoured bidder).

When these indicators pass a certain threshold of 'normality', they trigger automated alerts for further investigation by public authorities, shifting the focus from random sampling to evidence-based risk management.

The central vision: Common European Data Spaces (CEDS)

While AI provides the brainpower, the CEDS provide the infrastructure for collective intelligence. This is the central concept of the CEDAR vision: the project may serve as a bridge to the Public Procurement Data Space (PPDS) and other sectoral spaces like the European Health Data Space and the Green Deal Data Space.

In the past, an anti-corruption tool developed in one Member State was only as good as the local data it was trained on. CEDAR breaks this limitation. By connecting to the CEDS, the project enables a trans-European feedback loop:

- **Sharing aggregated results**: instead of sharing raw, sensitive data, Member States can share the 'aggregated results' of their analyses. If a specific pattern of fraud is detected in the medical procurement sector in Italy, the parameters of that 'indicator' can be shared across the CEDS.
- **Updating EU-wide indicators**: other nations can then update their local AI models with these new parameters. This creates a dynamic, evolving defence system where an attempt at corruption in one corner of the EU instantly makes the entire union more resilient.
- **Cross-sectoral enrichment**: by linking procurement data with the European **Health Data Space (EHDS)** or the **Green Deal Data Space**, CEDAR can monitor, for example, whether 'green procurement' targets are actually being met or if pharmaceutical spending is optimized across borders.

Progress towards greater transparency

Currently in its last year, the CEDAR project – a consortium of 31 partners including research institutions, tech leaders, and non-governmental organizations (NGOs) – is validating its technologies through high-impact pilots. Pilots in Italy, Slovenia and Ukraine focus on the health sector, analysing historical bidding data from hospitals and ministries to identify risk patterns in medical supply chains.

The project's progress is measured not just in lines of code, but in the creation of **interoperable connectors and AI tools** that allow any public organization to plug into the CEDS ecosystem and get deep analysis. By 2026, CEDAR aims to have successfully shared novel AI-driven indicators, high-quality procurement datasets and enriched different CEDS, providing a blueprint for the EU's 'Digital Decade 2030'.

Conclusion

The CEDAR project represents a fundamental shift in how we view the relationship between the state and data. By moving from reactive oversight to proactive, AI-driven management, the EU is setting a global standard for transparent governance. Through the use of CEDS, CEDAR ensures that public procurement is no longer a collection of isolated transactions, but a unified, intelligent ecosystem. In this new era, data can become a tool for building a more fair, competitive, and trustworthy European future.

cedar-heu-project.eu

CEDAR has received funding by the European Union's Horizon Europe research and innovation programme under grant agreement no. 101135577. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union. Neither the European Union nor the granting authority can be held responsible for them.

Since January 2024, the SPECTRUM project has been investigating a compute and data continuum to meet the needs of data-intensive science communities. In this article, Sergio Andreozzi and Xavier Salazar (both EGI Foundation) unveil SPECTRUM's conclusions and show how these impact the HiPEAC community.

A shared vision for the future of data-intensive science in Europe

SPECTRUM

June 2026 marks the close of SPECTRUM, a European-funded initiative to shape the future of data-intensive science in Europe, in close collaboration with relevant scientific communities – in particular the high-energy physics (HEP) and radio astronomy (RA) communities.

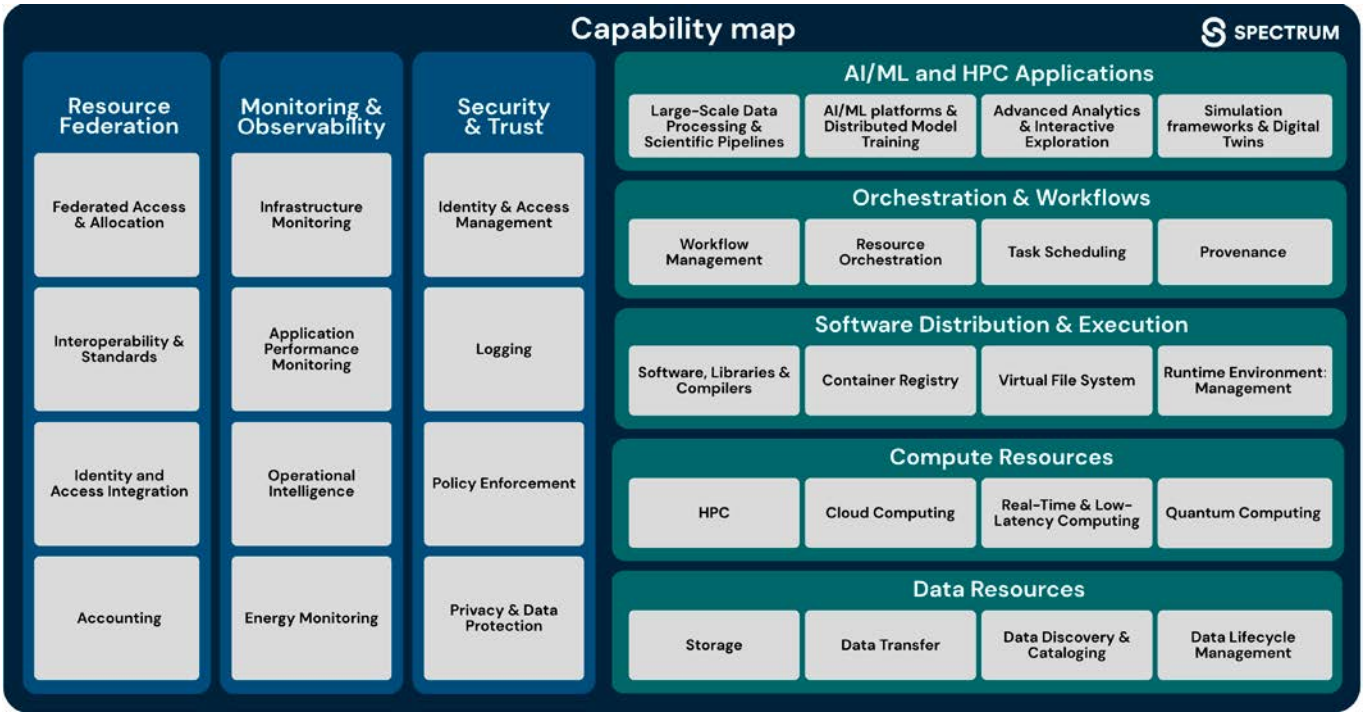
European research communities are preparing for decades of data growth and increasing computational complexity. Research infrastructures are generating unprecedented data volumes, and will increasingly rely on highly sophisticated workflows for simulation, analysis, and real-time processing.

Incremental improvements to existing infrastructures will not be sufficient to meet these requirements: there is a need for a coherent **compute and data continuum** that links high-performance computing, cloud platforms, edge devices, and large-scale data services in a way that is reliable and easy to use for scientific communities.

After two and a half years, SPECTRUM is now presenting its two main outcomes: **the Strategic Research, Innovation and Deployment Agenda (SRIDA)** and the **Technical Blueprint**. Both documents build upon prior work conducted within the project, including insights from the community of practice (SPECTRUMCoP), public surveys, interviews with relevant experts, use-case analyses, and assessments of the infrastructure landscape. The Technical Blueprint provides the architectural foundation, while the Strategic Research, Innovation and Deployment Agenda (SRIDA), provides the strategic framework for investment, governance, and implementation. Together, they chart the path toward a coherent European compute and data continuum for science.

SPECTRUM Technical Blueprint

The SPECTRUM Technical Blueprint sets out how Europe can move from today's fragmented landscape toward an integrated, long-term continuum. By clearly defining which **core capabilities** are needed, and identifying the most pressing **gaps**, it outlines a **practical path** for connecting compute, data, software, and security into a coherent infrastructure that is able to address the challenges for future data-intensive research. The approach is designed to be both technically realistic and aligned with the operational requirements of HEP, RA, and other scientific domains, providing clear and actionable recommendations for Europe's future e-infrastructure designs and research funding programmes.



SPECTRUM's Technical Blueprint sets out a pathway for scientific infrastructures

By addressing these priorities, the Technical Blueprint directly supports the European policy goals of digital sovereignty, economic and environmental sustainability, and strategic autonomy. It provides a practical pathway to infrastructures that not only serve flagship communities such as HEP and RA, but also strengthen Europe's position in the landscape of data-intensive science as a whole, fully exploiting Europe's scientific potential.

Main gaps and actions

The analysis identifies significant barriers to achieving a production-ready continuum and recommends coordinated European efforts to address them. Among recommendations of particular interest to the HiPEAC community are the following:

- Interfaces and standards must be harmonized, since interfaces between infrastructures are often inconsistent, making **interoperability** a persistent challenge.
- **Co-design** processes should be strengthened: currently, co-design between scientific communities and infrastructure providers is not systematic, which risks creating mismatches between what is built and what is needed.
- **Software portability and execution environments** need investment: the portability of software is still limited and applications are often unable to efficiently exploit emerging hardware.
- **Workflow orchestration** must evolve into a **cross-facility** capability with built-in resilience and provenance. Current workflow management tools are not capable of seamless execution across multiple sites, and provenance capture is inconsistently applied.
- A federated **trust and security framework** should be adopted, building on existing identity federations but extending them to new services and disciplines. The fragmentation of existing mechanisms slows down cross-border and cross-domain collaboration.
- **Artificial intelligence and machine learning (AL / ML)** should be integrated as a core capability in data-processing pipelines, from data acquisition to simulation and analysis. To ensure **long-term sustainability**, policies regarding the use and long-term management of software and hardware resources must be implemented. These should cover management of environmental impact of the full lifecycle of the infrastructure, operational funding, and workforce development.

SPECTRUM Strategic Research, Innovation and Deployment Agenda (SRIDA)

The Strategic Research, Innovation and Deployment Agenda (SRIDA) provides the framework for this transformation, presenting priorities that enable each stakeholder group to contribute to a coherent European response. Building on operational experience from federated infrastructures that already coordinate nationally funded resources at European scale, the SRIDA is the strategic complement to the Technical Blueprint, spanning key policy actions and investment areas.

Recommended actions

- **Scientific communities** should publish multi-year resource plans, participate in co-design processes with e-infrastructures, and identify high-impact use cases for AI / ML transition to production.
- **Service providers** should establish liaison mechanisms with thematic research infrastructures, deploy federated identity access management and standardized interfaces, and provide training enabling efficient resource use.
- **Policy makers** should ensure data-intensive science requirements inform EuroHPC / European Open Science Cloud (EOSC) priorities, enable multi-year allocation mechanisms, and fund sustained software development, data management, and career frameworks. They should also ensure that the environmental impact of research infrastructures are within recommended limits.

Investment areas

Key areas for investment include computing infrastructures, data infrastructures, software and tools, governance and coordination, and human capital. Open science and environmental sustainability cut across all areas. Coordination costs, sustained software development, and long-term preservation of data using FAIR best practices, while often underestimated, are essential for realizing value from capital investments. Effective investment requires coordination across European and national programmes, and research infrastructure operational funding.

What's next?

Find out more about outcomes from the SPECTRUM project and read more about our future plans on the SPECTRUM project website and on our LinkedIn page – see 'Further information', below.

SPECTRUM would like to thank all the experts who contributed to the creation of these documents.

FURTHER INFORMATION:

SPECTRUM website spectrumproject.eu
SPECTRUM on LinkedIn linkedin.com/company/spectrum-project-eu
SPECTRUM Technical Blueprint spectrumproject.eu/technical-blueprint
SPECTRUM SRIDA spectrumproject.eu/srida

HAR.S.H. project: HARDware-aware extreme-scale Similarity search

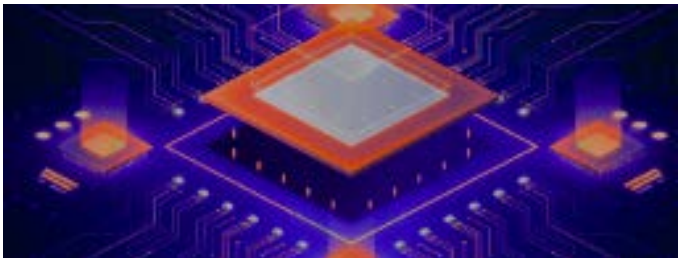
Modern scientific and industrial applications generate massive volumes of time series and multi-modal data, including text, images and video. Analysing such data at scale requires efficient, hardware-conscious solutions able to harness the full power of modern computational platforms. The HAR.S.H. project has been tackling this challenge by developing algorithms, data structures, and mechanisms for efficient similarity search over large-scale time series collections, considering the architectural characteristics of modern and emerging hardware technologies. The project will produce a strong, reusable software library, enabling its adoption by a wide range of real-world applications.

The HAR.S.H. project is funded by the Ministry of Education, Religious Affairs and Sports, under the Greece 2.0 National Recovery and Resilience Plan. The goal of HAR.S.H is to bring together leading researchers and companies to combine efforts in order to address the main challenges in large-scale data processing. Specifically, it will have significant impact in the following directions:

1. **Hardware-aware similarity search**
   Design scalable algorithms and data structures that enable parallel and distributed similarity search for vast collections of time series data.
2. **Leveraging emerging hardware**
   Investigate and exploit the effects of hardware platforms, including memory, synchronization, and communication technologies, on boosting performance and scalability.
3. **Multi-modal data integration**
   Develop embedding and indexing techniques that allow for deep learning-based analysis of multimodal data, including text, images, and video.

HAR.S.H has demonstrated its capabilities with real-life datasets from demanding application areas and industrial benchmarks, including three application pilots from the enterprises partners. Specifically, HAR.S.H. has been focusing on the following pilot applications:

- **Pilot application 1**: Search for similar documents and files in large document databases.
- **Pilot application 2**: Photo analysis for improving travel profiles. The main goal is to maximize visitor satisfaction with a tourist destination.
- **Pilot application 3**: Public opinion analysis application by monitoring and categorizing content from social media.



Impact

By combining state-of-the-art algorithmic design with hardware-aware optimizations, HAR.S.H. is pushing the limits of scalable time series and multi-modal data analysis. Its technological outputs will be vital for applications in areas such as smart tourism, digital heritage, sentiment analysis, and large-scale information retrieval. The project also contributes to the European data and artificial intelligence (AI) strategy, offering a practical foundation for next-generation data-intensive systems.

FURTHER INFORMATION:

PROJECT NAME: HARSH: Hardware-Aware Extreme-Scale Similarity Search
START / END DATE: 15/04/2025 – 31/06/2026
KEY THEMES: similarity search, time series, hardware-aware algorithms, deep learning, multi-modal data
COORDINATOR: ELKE-UOC: Special Account for Research Funds - University of Crete
COLLABORATING ORGANIZATIONS: AMS (Greece), MCBS (Greece), NEUROLINGO (Greece)
CONTACT: Professor Panagiota Fatourou, University of Crete and FORTH-ICS
faturu@csd.uoc.gr
harsh-project.eu/en
www.linkedin.com/company/harsh-project]
facebook.com/harshproject

HAR.S.H. (grant agreement no. ΥΠ3ΤΑ-0560901) is carried out within the framework of the National Recovery and Resilience Plan 'Greece 2.0' with funding from the European Union – NextGenerationEU.

Harnessing data to build trust in European governance: The CEDAR project

In recent years, Europe has been facing unprecedented crises on multiple fronts, including health, migration, the economy, climate, energy, and politics. This has led to a sharp increase in emergency public spending and the relaxation of due diligence checks, which, in turn, has resulted in a rise in corruption and fraudulent activities, with significant negative impacts on the European economy, society, environment, and democracy.

Despite emerging technology's potential to become a powerful tool in the fight against corruption and fraud, the public sector has been slow to adopt digitalization, resulting in data not being shared, harmonized, or properly analysed, making evidence-based decision-making almost impossible. Governments are slowly adopting new approaches to ensure more data-driven, transparent, and accountable public governance, but several fundamental data-related issues remain unresolved.

With a team of nine excellent research institutions and universities, 12 technology, business, and standards, developing companies, seven public end users, and three domain-relevant, industry-exposed non-governmental organizations (NGOs), CEDAR is:

1. Identifying, collecting, fusing, harmonizing, and protecting complex data sources to generate and share more than 10 **high-quality, high-value datasets** relevant for more transparent and accountable public governance in Europe.
2. Developing interoperable and secure **connectors and application programming interfaces (APIs)** to utilize and enrich at least six Common European Data Spaces.
3. Developing innovative and scalable technologies for effective **big data management and machine learning (ML) operations.**
4. Delivering robust **big data analytics and ML** to facilitate **human-centric and evidence-based decision-making** in public administration.
5. **Validating the new datasets and technologies** (to technology readiness level 5, or TRL5) in the context of fighting corruption, thus aligning with the European Union's strategic priorities of digitalization, the economy, and democracy.
6. Actively **promoting results across Europe** to ensure their **adoption and longevity**, and to generate positive, direct, tangible, and immediate impacts.

A suite of deliverables and other resources – including videos and podcasts – can be accessed via the project's website.

PROJECT NAME: CEDAR - Common European Data Spaces and Robust AI for Transparent Public Governance
START / END DATE: 1 January 2024 – 31 December 2026
KEY THEMES: Common European Data Spaces, digitalization of public administrations, artificial intelligence, transparency
PARTNERS: **Greece**: Centre for Research and Technology Hellas (coordinator), DBC diadikasia; **Belgium**: KU Leuven, The Lisbon Council for Economic Competitiveness; **Cyprus**: Ubitech; **France**: CEA (French Alternative Energies and Atomic Energy Commission); **Germany**: Brandenburg Institute for Society and Security (BIGS), Kiel University (CAU), Transparency Germany; **Italy**: CENTAI, Insiel, Autonomous Region of Friuli Venezia Giulia, ANCE Friuli Venezia Giulia industrial association, Arpa FVG; **Luxembourg**: Netcompany; **Netherlands**: Arthur's Legal; **Slovenia**: Snep, Institute for Corporative Security Services, Engineering, Slovenian Ministry of the Interior, Slovenian Ministry of Health, Slovenian Ministry of Digital Transformation, Celje General Hospital; **Spain**: Vicomtech, Trebe Language Technologies, Universidad Politécnica de Madrid; **Ukraine**: YouControl, Artellence, State Agency for Restoration and Development of Infrastructure of Ukraine, SK Security, Safe Ukraine 2030
BUDGET: € 9 million

cedar-heu-project.eu
linkedin.com/company/cedar-eu
@cedar_eu
@cedar-eu.bsky.social
youtube.com/@CEDAR_EU

Continuing our series on innovative deep-tech companies in Europe, we learn about Logiicdev, a Graz-based platform developer with a majority female team, who are developing edge AI and reconfigurable solutions for robotics and manufacturing markets.

Logiicdev

Smart, reconfigurable platforms for the next-generation IoT

COMPANY: Logiicdev GmbH
MAIN BUSINESS: integration of artificial intelligence (AI) and digitization at the hardware, software and application layers, reconfigurable computing (field-programmable gate arrays – FPGAs), internet of things (IoT)
LOCATION: Graz, Austria
WEBSITE: logiicdev.com

Based in Austria, Logiicdev GmbH is a high-tech small / medium enterprise (SME) delivering edge AI and FPGA-based solutions for robotics and manufacturing markets. The company's research-driven approach combines computer vision, digital-twin technology, and real-time processing into compact, low SWaP-C (size, weight, power and cost) architectures, empowering industries to digitize and optimize operations without compromise on reliability or performance.

Logiicdev is active as industrial solution enabler, research partner, and intellectual property (IP) developer, driving the European Union (EU)'s vision by contributing to different working groups, joint undertakings and clusters. The company's green-energy management solutions also align with the EU's ambition to achieve 40% renewable energy by 2030. Logiicdev is researching advanced technologies for 5G connectivity, monitoring, integrating, and optimizing renewable resources across grids, communities, and industries. Its edge AI and IoT platforms enable real-time management, flexibility services, and smart grid balancing, while optimizing energy use in buildings, industry, and mobility. These platforms can also be used for next-generation industrial automation.

Logiicdev is also developing FPGA-based solutions for post-quantum cryptography (PQC), which is critical for protecting digital systems in the era of quantum computing. FPGA-based PQC combines quantum-resistant algorithms with reconfigurable hardware to deliver a scalable and future-proof security foundation.

Unlike fixed application-specific integrated circuits (ASICs) or software-only solutions, FPGAs enable customizable architectures that evolve with emerging standards and threats. In addition, FPGA implementations are optimized for low power consumption, reduced logic utilization, and fewer clock cycles, making them well suited for embedded devices, edge computing, and high-throughput secure systems. Native parallelism and pipelining accelerate lattice-based and hash-based cryptography while ensuring low latency and deterministic performance, while modular design enables flexible scaling of security levels, key sizes, and performance without hardware changes.

Among the company's products are:

- **KiiGen**: an industrial edge AI platform for smart automation (see pp.32–33 for a full exploration of this platform)
- **V.nodTM**: a drone-energy management platform for verified, predictable and reusable battery operations
- Tailored FPGA cores

The team has exceptionally high female representation compared to other organizations in the field: its composition is two-thirds women to one-third men. Logiicdev is currently recruiting and offers a range of options for students and recent graduates, including internships, part-time roles for working students, and PhD opportunities.

Logiicdev is based in the picturesque city of Graz, Austria
Photo: Fabio Lotti | stock.adobe.com

Based in Barcelona, Mosaic Factor offers a range of services including data analytics, machine learning, optimization models, digital twins and large language models (LLMs), with trustworthy and explainable artificial intelligence (AI) as a key concern.

Mosaic Factor

Data-driven solutions tailored to real-world operational environments

COMPANY: Mosaic Factor
MAIN BUSINESS: applied artificial Intelligence, data engineering, and advanced analytics, with a focus on operational AI systems, trustworthy AI, and data-driven products
LOCATION: Barcelona, Spain
WEBSITE: mosaicfactor.com

Mosaic Factor is a boutique AI and big data engineering firm based in Barcelona, focused on designing and deploying data-driven solutions for real-world operational environments. Founded in 2016 and celebrating its tenth anniversary, the company has delivered more than forty innovation projects and collaborated with over thirty-five clients and partners across European ecosystems.

Its work spans the full data and AI lifecycle, including data analytics, machine learning, optimization models, and AI-based solutions such as digital twins and domain-specific LLMs. These capabilities are applied across sectors including mobility, logistics, automotive, manufacturing, and healthcare, often in contexts where AI systems must operate under real constraints and directly support business-critical decisions.

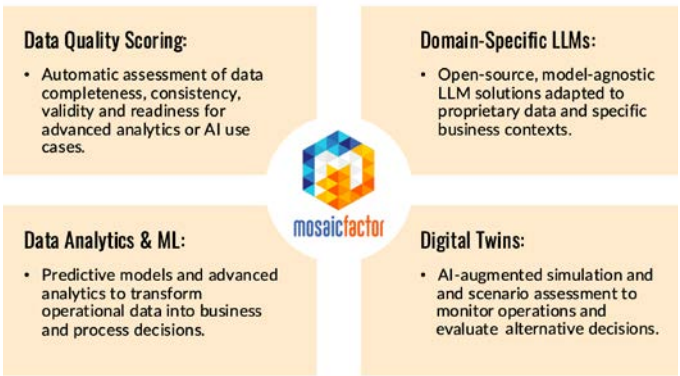
A defining characteristic of Mosaic Factor's approach is its focus on operational deployment rather than experimental prototypes. The company develops AI systems that are designed to function in dynamic environments, integrate with existing processes, and deliver measurable impact. Typical applications include digital twins, light domain-specific LLMs, supply chain optimization, and quality analytics, where model outputs need to be reliable, interpretable, and continuously monitored.

Within this context, the company has developed strong expertise in trustworthy AI, with a particular emphasis on explainable AI (XAI). From an engineering perspective, explainability is not treated as an isolated feature, but as an integral part of building deployable AI systems. The company addresses a key challenge in industrial AI adoption: the gap between high-performing models and their effective use in decision-making. In practice, even accurate models may fail if users lack sufficient visibility into their behaviour to trust, validate, or challenge their outputs.

To address this, Mosaic Factor adopts a structured, decision-centric approach to explainability. Rather than starting from specific techniques, it begins with the decision context: the business objective, the cost of incorrect predictions, and the needs of different stakeholders. Based on this, the company designs tailored explainability strategies combining interpretable modelling approaches, post-hoc methods such as feature attribution and sensitivity analysis, and process-level transparency mechanisms. These are integrated into broader governance frameworks, including validation pipelines, monitoring, and human-in-the-loop supervision.

Importantly, explainability is not considered separate from the broader AI stack. It builds on robust data foundations, careful model design, and a clear alignment with business objectives. This integrated perspective enables Mosaic Factor to deliver AI systems that are not only technically sound, but also usable and trustworthy in operational settings.

Looking ahead, Mosaic Factor aims to further advance the integration of scalable, trustworthy AI in complex, data-intensive environments. Its objective is to support organizations in moving from experimental AI adoption to reliable, interpretable, and accountable systems capable of supporting high-impact decisions in real-world conditions.



In February, the artificial intelligence (AI) research company Anthropic refused to allow certain uses of its Claude models by the US Department of Defense (DoD). In this opinion piece, Adam Mackay (Head of AI, QA-Systems), argues that, rather than only being a political crisis, this also exposes a safety engineering dispute, and a concrete compliance risk for European companies building high-risk AI systems for global markets.

The safety requirement the US refused

In February 2026, the US DoD told Anthropic to open its Claude models for any 'lawful purpose', including fully autonomous weapons or mass surveillance. Anthropic refused. In response, the company was tagged a 'supply chain risk to national security'. It was reportedly the first time this label had been applied to a US company, and the designation followed closely on the collapse of contract negotiations. Within hours, OpenAI announced it had signed a deal with the DoD on broadly similar terms.

The press reported the Anthropic / DoD dispute as a political story. For anyone who works with functional safety standards, it also has a clear safety-engineering dimension.

In safety-critical systems, a safety requirement is not a policy wish or a pledge of good intent. It is a testable constraint on system behaviour, checked and proved. The full architecture of DO-178C, ISO 26262, and IEC 61508 rests on this idea. You do not trust the pilot not to exceed structural limits. You design the flight envelope so that certain states are unreachable.

Anthropic asked for the same structure: clear contract terms that state which use cases are allowed and which are banned. The parallel is imperfect (large language model (LLM) use cases are emergent, not physically constrained), but the engineering principle is the same. The DoD said no. It wanted the right to use the models for 'all lawful purposes' and would not accept hard limits on any class of use. In safety-engineering terms, that is a refusal to accept a safety requirement. The US government then tagged the company that insisted on one as a national security risk.

The counter-argument is that law alone provides sufficient constraint. It does not. As legal scholar Jessica Tillipman of George Washington University noted, the OpenAI contract does not give the company a 'free-standing right to prohibit otherwise-lawful government use'. The contract defers to whatever the law says at the time rather than setting its own hard limits. Anyone who has written a safety case knows the gap this creates. A clause that says 'the system shall not perform function X' is a hard interlock. A clause that says 'the system shall comply with applicable law regarding function X' is a procedural control. Every safety standard ranks the hard interlock above the procedural control.

The EU AI Act (Regulation 2024/1689) points in the same direction: towards documented, enforceable constraints on how AI systems are used and controlled. Article 5 prohibits specific AI practices, including real-time remote biometric identification in public spaces and certain forms of social scoring. Articles 9 and 14 require documented risk management across the full system lifecycle. They also require high-risk AI systems to support effective human oversight, including the ability to intervene or override where appropriate. These are legal, enforceable requirements.

Consider a European company building high-risk AI systems, or components likely to be integrated into them, and selling into the US market. The Anthropic case shows that a US government buyer can blacklist a vendor for insisting on the very safety constraints that EU law requires. This creates a compliance and market-access risk for any European company whose activity falls under the AI Act: it may need to preserve hard safety constraints to satisfy European obligations, while facing pressure from a US government buyer to remove or weaken those constraints. Developing different configurations for different jurisdictions may reduce some legal risk, but it does not remove the underlying procurement risk if a buyer rejects the vendor's right to impose hard limits at all.

The conflict raises a direct question for the HiPEAC community: where in the computing stack does safety enforcement actually live? If a government buyer can refuse contract terms, and if policy can shift between one administration and the next, then enforcement cannot rest on contracts or rules alone. It has to be structural, i.e. built into the system itself.

Safety-critical engineering already applies this principle at the highest integrity levels. ISO 26262 Part 5 requires hardware safety mechanisms for ASIL C and D systems. Software on its own cannot reach the needed diagnostic coverage. The standard therefore demands a separate, independent enforcement path. IEC 61508, and its sector application IEC 61511, takes the same line: safety-instrumented systems must be separate from and independent of the process they monitor. Across every domain where failure is catastrophic, the rule is the same: safety enforcement must be structurally independent of the system it constrains.

If a foreign government buyer can turn a European company's compliance duties into a market-access risk, the question is not whether structural enforcement is needed but where it belongs in the technical architecture. The HiPEAC community designs processor architectures, compiler toolchains, runtime checks, and hardware-software co-design frameworks. It has the technical depth to answer that question. Without that answer, European AI companies may find it harder to compete globally while still holding the safety line that European regulation draws.

FURTHER READING:

EU AI Act (Regulation 2024/1689)
eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

'OpenAI's "compromise" with the Pentagon is what Anthropic feared', MIT Technology Review, 2 March 2026
bit.ly/MIT_TR_Pentagon_Anthropic

Jessica Tillipman, 'What Rights Do AI Companies Have in Government Contracts?', Nextgov/FCW, March 2026
bit.ly/Jessica_Tillipman_2026

'OpenAI strikes deal with Pentagon, hours after rival Anthropic was blacklisted by Trump', CNBC, 27 February 2026
cnb.cx/4nHDyLx

'Trump admin blacklists Anthropic as AI firm refuses Pentagon demands', CNBC, 27 February 2026
cnb.cx/4rO6QJS

Expand your research horizons with the HiPEAC Student Challenge

The HiPEAC Student Challenge is an opportunity to test your programming and / or computer architecture skills, tackle real-life problems, learn-by-doing research methodologies, and connect with peers.

Once again, the HiPEAC Student Challenge will be organized at the HiPEAC conference (18–20 January 2027, Glasgow). There will be different kinds of challenge to choose from, ranging from open challenges to reproducing a scientific paper or contributing to the RISC-V ecosystem. Interested teams will be asked to register their interest by 10 September.

Up to five student teams will be selected (with each team comprising a maximum of five students). From these selected teams, up to 12 students in total will be invited to attend the HiPEAC conference and present their work. Registration fees, travel and accommodation expenses will be covered by the organizers.

Further details will be available in due course – check the HiPEAC website and social media for updates.

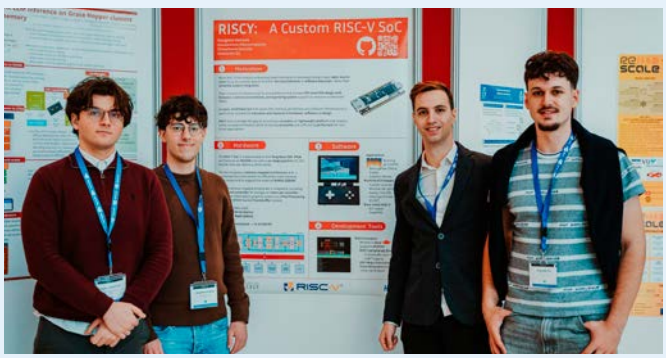
hipeac.net
hipeac.net/linkedin

HiPEAC 2026 Student Challenge débuts RISC-V category to great effect

At HiPEAC 2026, the HiPEAC Student Challenge included a new category. In addition to the open challenge and the reproducibility challenge, for the first time students were invited to contribute to the RISC-V ecosystem – and they took up the challenge in style. Students at the University of Thessaly's Computer Systems Lab came up with RISCY (a custom RISC-V field-programmable gate array (FPGA) system on chip (SoC) – and a small game console (running DOOM). As well as presenting in the Student Challenge session, team representative Panos Nanousis also demonstrated the project at the RISC-V for high-performance computing (HPC) workshop, where it was one of the most popular presentations.

Meanwhile, Amer Mujalo presented work by his team at the University of Sarajevo on VHDL code generation, while Panagiotis Gialamas and George Triantafyllou (ICCS / NTUA) presented their smoke-detection tool. Mert Karahan (Izmir Institute of Technology) and Simone Staccone (Tor Vergata University, Rome) presented their teams' reproduction of scientific papers.

HiPEAC would like to congratulate all those who took part and to thank the organizers, Marisa Gil, Raffaella Folgieri, Işıl Öz, Teresa Cervero García, Georgios Goumas, and Chris Fensch.



In this edition of our regular series showcasing early career researchers' work, we learn how Aleksandra Poreba developed novel, efficient ways to accelerate data-intensive workloads from the ATLAS Experiment at CERN using graphics processing units (GPU).

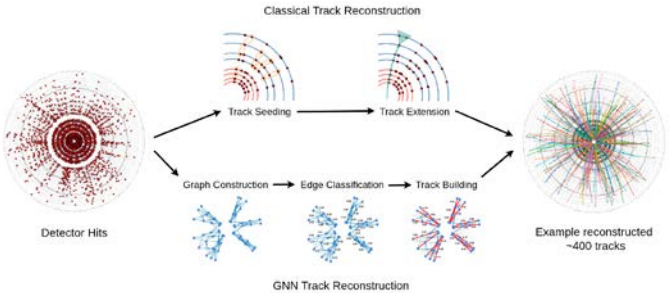
Three-minute thesis

NAME: Aleksandra Poreba
RESEARCH CENTRES: CERN and Heidelberg University (HAWAII Lab, ZITI)
SUPERVISORS: Holger Fröning (Heidelberg University), Brian Petersen (CERN)
THESIS TITLE: How to connect the dots very fast? High-performance heterogeneous particle track reconstruction for the ATLAS Phase-II High Level Trigger

High-energy physics experiments such as the ATLAS Experiment at CERN pose unique data-processing challenges. The trigger system, responsible for filtering detector data from particle collisions in real time, must operate under strict latency constraints while maintaining extremely high accuracy. Any inefficiency or loss in performance may result in the loss of valuable data essential for particle physics research.

One of the most computationally demanding tasks is particle track reconstruction, which involves identifying the helical trajectories of charged particles from a cloud of points formed from detector measurements ('hits'). A single collision event may contain approximately 350,000 hits corresponding to around 1,500 particle tracks. In the ATLAS trigger system, track reconstruction algorithms must execute at rates of 200 kHz on a dedicated computing farm while maintaining high reconstruction accuracy. My PhD research focused on demonstrating and optimizing GPU-accelerated approaches to particle track reconstruction, including both classical and machine-learning-based methods.

The first stage of the classical tracking pipeline is track seeding, which identifies three starting points for particle trajectories ('seeds'). The task is well-suited to GPU acceleration because it is highly parallelizable and limited to local detector regions. The main computational challenge is reducing the number of incorrect seeds, as each false candidate triggers the computationally expensive track extension stage. My work introduced a novel seed-confirmation algorithm that, together with improved GPU resource allocation strategies, significantly improved reconstruction accuracy and computing performance.



Tracking pipelines

The machine-learning-based tracking approach uses graph neural networks (GNNs), where detector hits are represented as graph nodes, and track segment candidates as edges. To classify the graph edges, information within the graph is exchanged via a message passing mechanism. The main computational challenge is driven by the scale of the input graphs, which may contain up to three million edges per collision event.

Processing the full graph at once leads to a substantial memory footprint, limiting deployment on memory-constrained GPUs. Conventional graph-partitioning techniques can reduce memory usage, but they require careful handling of overlaps between partitions to reconstruct all shapes of particle tracks. Instead, I developed a partial message processing approach operating on temporary fixed-size structures. This simple and lightweight strategy enabled parallel processing and efficient deployment on resource-constrained GPU systems while preserving tracking performance.

To accelerate GNN inference, I investigated model-compression techniques designed to accelerate GPU matrix-multiplication operations, including structured pruning. This provides a unique advantage, not only removing unnecessary computations but also scaling the model size. The proposed compression strategy was proven to successfully accelerate several types of GNNs, including interaction graph neural networks (IGNNs), graph convolutional networks (GCNs) and graph attention networks (GATs).

Conclusion

During my PhD research, I optimized and demonstrated the application of GPU-accelerated algorithms for low-latency, high-accuracy environments such as the ATLAS Trigger System. Both conventional and machine-learning-based tracking approaches present distinct challenges that must be addressed before deployment in real-time systems. The algorithms were evaluated across multiple GPU architectures, focusing on their differing hardware capabilities and performance characteristics.

SPONSORSHIP OPPORTUNITIES

HiPEAC conference

18 - 20 January 2027

Glasgow

Europe's largest computing systems research event (500-700 participants, 80+ companies)

Expand your business network and meet new clients

Excellent recruitment opportunity: HiPEAC Jobs Fair and STEM Student Day

Media coverage and enhanced visibility options

Tailored sponsorship plans

	PLATINUM €10,000	GOLD €5,000	SILVER from €2,500	BRONZE €1,500
Logo on HiPEAC website and communications	✓	✓	✓	✓
Conference passes	6	3	2	1
Industry session presentation	30 min	20 min	10 min	✗
Industry exhibition booth	✓	✓	✓	✓
Privileged booth location and customized options	✓	✓	✗	✗
Eligibility to sponsor specific activities	✓	✓	✗	✗
HiPEAC Jobs support	✓	✓	✗	✗
Additional promotional opportunities	✓	✓	✗	✗

sponsorship@hipeac.net

Funded by the European Union

bit.ly/HiPEAC27_sponsorship