The Serbian AI Factory Antenna (SAIFA): Bridging High-Performance Computing with Energy and Environmental Sustainability

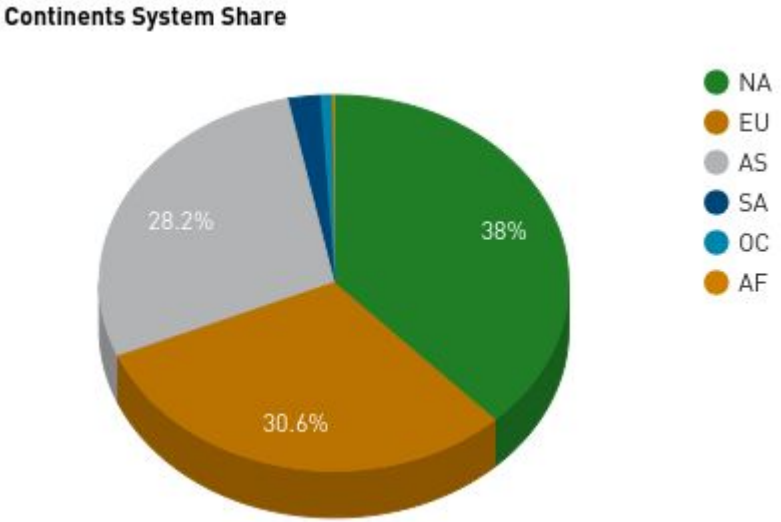
Miloš Ivanović, PhD, prof.
Faculty of Science, University of Kragujevac, Serbia
Faculty of Information Studies, Novo mesto, Slovenia

Belgrade, June 2026.

This project has received funding from the European High Performance Computing Joint Undertaking (EuroHPC JU) under grant agreement No 101263280.

EuroHPC Machines

- JUPITER (930 PFlops)
- LUMI (539 PFlops)
- LEONARDO (315 PFlops)
- MARENOSTRUM 5 (314 PFlops)
- MELUXINA (18 PFlops)
- KAROLINA (13 PFlops)
- VEGA (10 PFlops)
- DISCOVERER (6 PFlops)
- DEUCALION (10 PFlops)
- DAEDALUS (115 PFlops)
- ARRHENIUS (60 PFlops)



New AI Trends

- **A Paradigm Shift in Computing:** The rapid evolution of AI has necessitated a fundamental shift in computational infrastructure, bringing High-Performance Computing (HPC) into sharp focus for the entire global economy.
- **From Elite Research to Essential Foundation:** Historically, solving complex optimization and computational problems was the exclusive privilege of large enterprises and research institutions that could afford specialized expertise and massive infrastructure. Today, this is no longer the case as AI moves from experimental labs into every sector of society.
- **AI as a "Public Utility"**: AI services are no longer a luxury; they have become a critical utility similar to electricity or water, essential for powering modern healthcare, energy management, and environmental sustainability.

EuroHPC – Joint Undertaking

- Joint Undertaking of EU, Participating States & Private Partners
- The initiative acts as **a bridge between HPC and Artificial Intelligence**, supporting sovereign compute, large-scale training, and the development of foundation models
- Develop, deploy, extend & maintain in Europe a world-leading HPC supercomputing, quantum computing, service & data infrastructure ecosystem
- The overall EU and Member State investment in supercomputing and AI Factories through the EuroHPC Joint Undertaking **will reach €10 billion for the 2021–2027 period**.
- **The Critical Role of AI Factories**: This transition is why AI Factories are so vital - they act as integrated ecosystems that bridge the gap between world-leading supercomputing resources and the practical, real-time needs of local industries, startups, and public administrations.

SERBIAN ARTIFICIAL INTELLIGENCE FACTORY ANTENNA

EC Announcement - AI Factories

European Commission ✔
22 October at 09:24 · 🌐

The EU's AI Continent ambition is officially in full gear. 😎

Six new AI Factories just launched in 🇨🇿🇱🇹🇳🇱🇷🇴🇪🇸🇵🇱, bringing the total to 19 across Europe. These hubs give startups and companies direct access to supercomputers to power the next wave of cutting-edge AI.

Plus, 13 AI Factory Antennas are now live in 🇧🇪🇨🇾🇭🇺🇮🇪🇱🇻🇲🇹🇸🇰🇮🇸🇲🇩🇨🇭🇬🇧🇲🇰🇷🇸, connecting local innovators to top-tier tech and expert support.

Why does this matter?

✅ For startups and companies: massive computing power and expert support for faster, smarter AI products.
✅ For citizens: better public services, faster healthcare, and stronger digital security.
✅ For everyone: artificial intelligence built on strict ethical standards to be trustworthy, human-centric, and made in Europe.

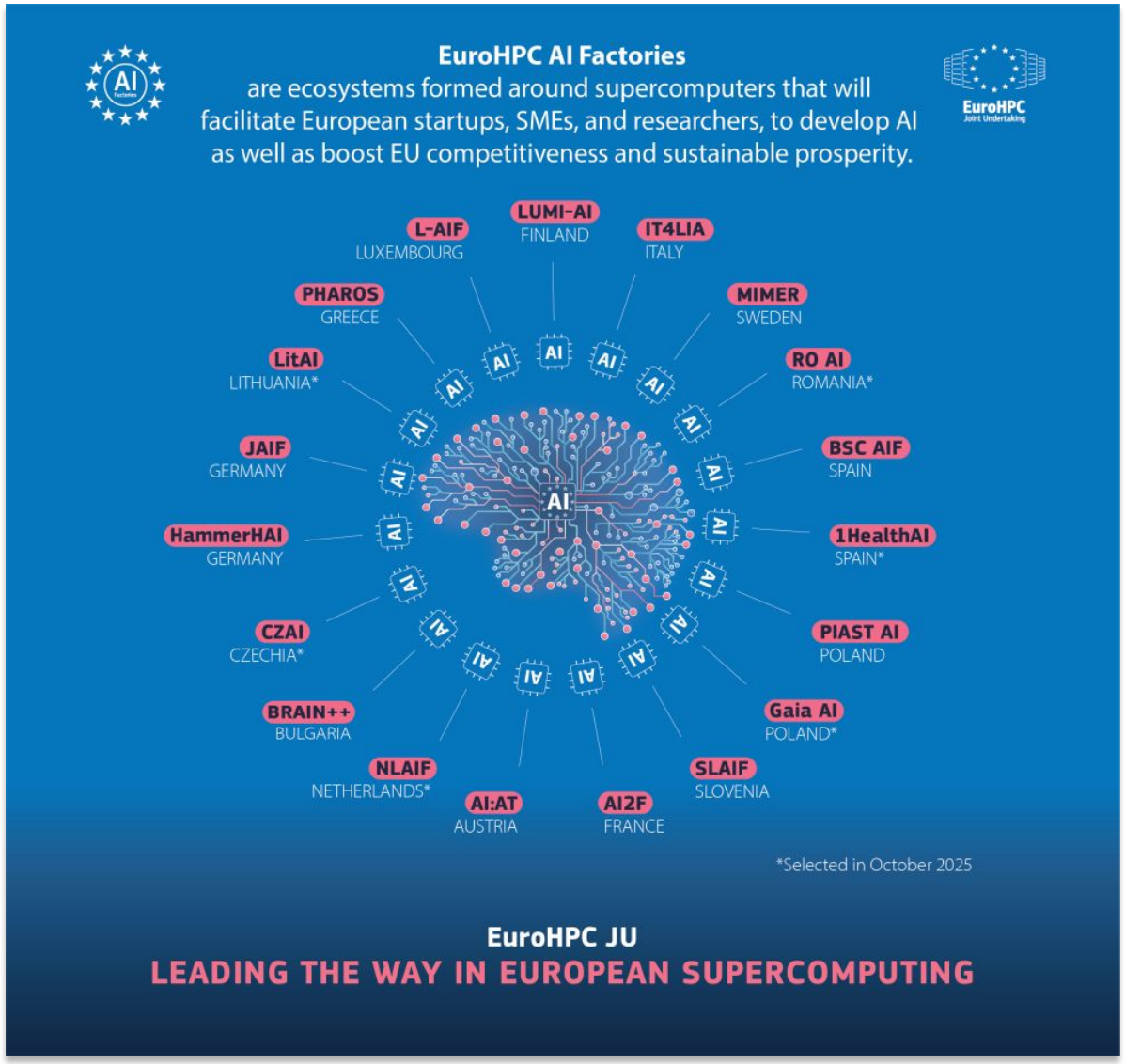
Time for our innovators to build the future, on Europe's strong foundations! 🚀

The total committed investment by the EU and participating EuroHPC countries for the AI Factories and Antennas initiative is over €2.6 billion.

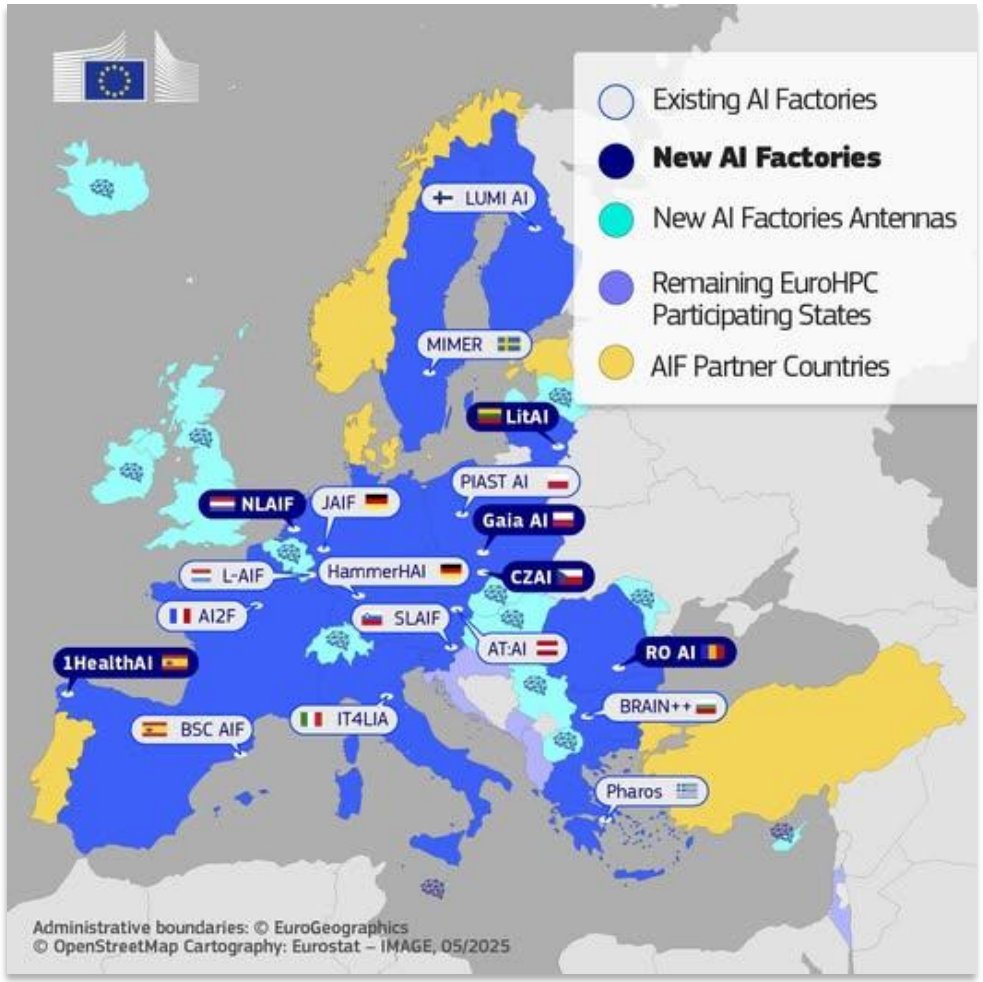
SERBIAN ARTIFICIAL INTELLIGENCE FACTORY ANTENNA

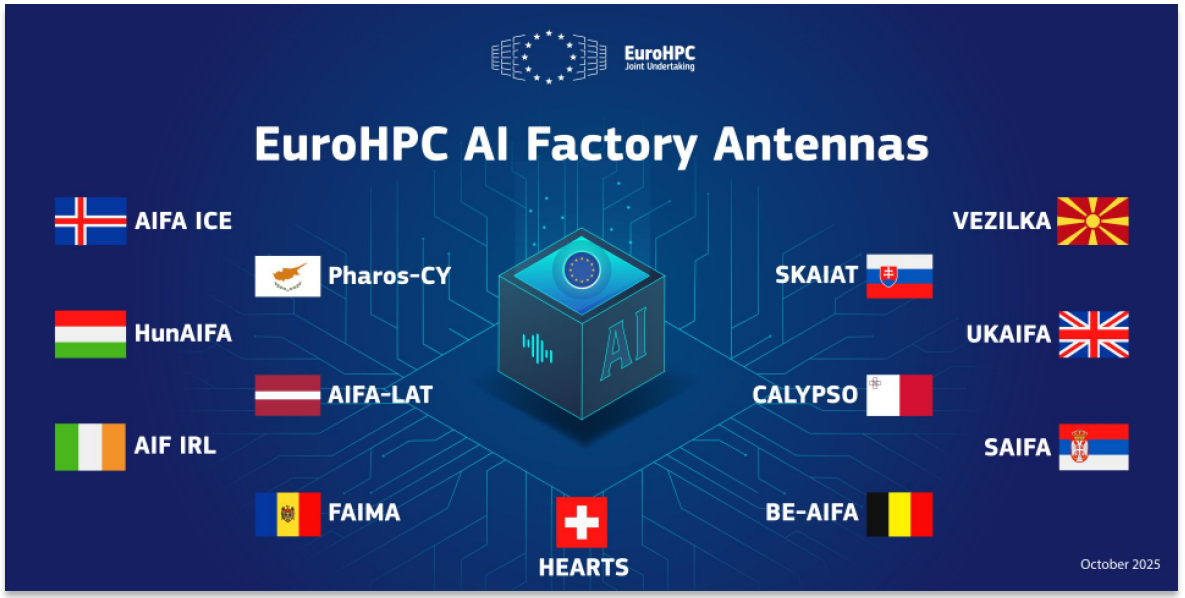
EuroHPC AI Factories

- European AI compute & data ecosystem under EuroHPC JU
- Bridges HPC + AI: sovereign compute, large-scale training, foundation models
- Integrates supercomputers, cloud, and edge capacities across Member States
- Supports AI R&D, startups, industry, public sector & science
- Common European AI stack: compute, storage, data lakes, tools
- Access through EuroHPC mechanisms: peer-reviewed, industry calls, innovation schemes



AI Factory (Antennas)





SAIFA Consortium - 8 partners

UNIVERSITY OF BELGRADE SCHOOL OF ELECTRICAL ENGINEERING

Lead partner

Institute of Physics Belgrade

Institute Mihajlo Pupin

FACULTY OF SCIENCE 1972 KRAGUJEVAC

Faculty of Science, University of Kragujevac

THE INSTITUTE FOR ARTIFICIAL INTELLIGENCE RESEARCH AND DEVELOPMENT OF SERBIA

Institute for Artificial Intelligence Research and Development of Serbia

ИМГГИ IMGGE

Institute of Molecular Genetics and Genetic Engineering, University of Belgrade

FEFA FAKULTET

FEFA / Uni Metropolitan

e Srbija

Office for Information Technologies and eGovernment

SAIFA SERBIAN ARTIFICIAL INTELLIGENCE FACTORY ANTENNA

SAIFA Computing Resources

Currently Available

- **National AI Platform (First Generation):** Features four NVIDIA DGX-A100 nodes delivering up to 5 petaFLOPS of AI compute power. It is optimized for deep learning and large-scale data processing with 120TB of high-bandwidth storage.
- **The PARADOX Supercomputer (PARADOX-IV)** at the Institute of Physics Belgrade consists of 106 working nodes powered by 1,696 Intel Xeon CPU cores and 106 NVIDIA Tesla M2090 GPU accelerators. It delivers a peak computing performance of 105.8 TFlops, supported by 3.3 TB of system memory and 100 TB of high-speed Lustre storage. Connected via high-throughput InfiniBand networking, this architecture is optimized for running complex numerical simulations, processing CERN data, and driving regional scientific research.

Upcoming Resources

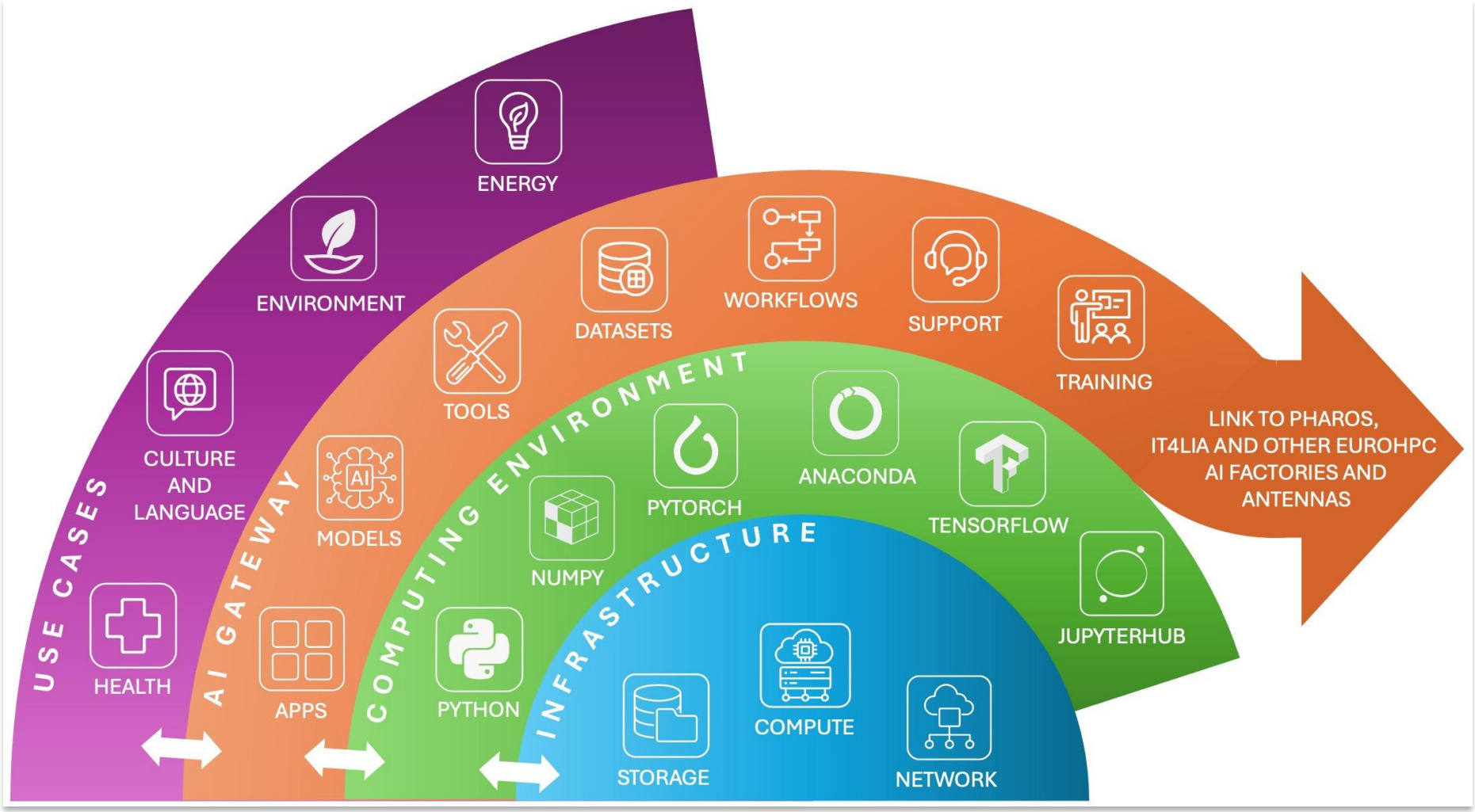
- **DGX-H200 Expansion:** The platform is being extended with a cluster of six DGX-H200 systems, adding 32 PFLOPs of compute power.
- **BullSequana XH3000:** This is the most advanced AI/HPC system in the SEE region, comprising 160 nodes powered by 640 NVIDIA GH200 Grace Hopper Superchips. It delivers over 22.5 PFLOPs of peak performance and is designed for exascale workloads like foundation model pre-training and digital twin creation.

Cross-border resources

- SAIFA provides a link to European AI Factories, such as Pharos (Greece) and IT4LIA (Italy), for final model validation and performance benchmarking.

SERBIAN ARTIFICIAL INTELLIGENCE FACTORY ANTENNA

SAIFA Architecture



SAIFA Services

Horizontal services

Capabilities that assist all users and application domains, regardless of the sector or use case

Vertical services

Solutions tailored to specific sectors, addressing the unique needs, data features, and operational challenges of particular domains

Sectors of Interest - Use Cases

Culture and language

- Hate speech detection model for Serbian language
- AI-driven workflow for constructing knowledge graphs
- Digital twin for studying collective social phenomena
- LLM-powered prediction of homologous proteins in non-human organisms

Health

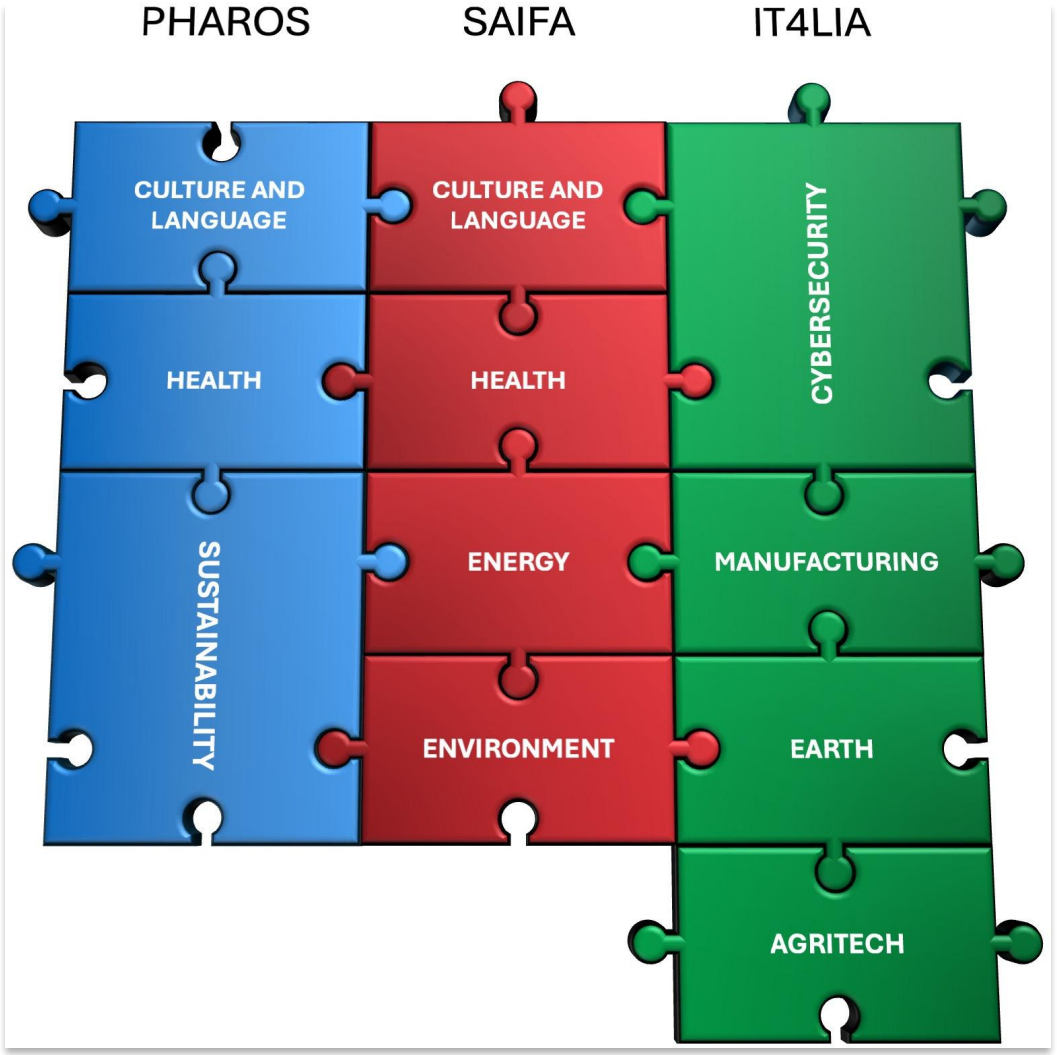
- Diagnosis Prediction
- Therapy Analysis
- Patient Assistance
- Robustness Challenges

Energy

- Energy generation and demand data analytics
- Process optimization based on surrogate models

Environment

- Flood prevention



Hub for AI Innovation

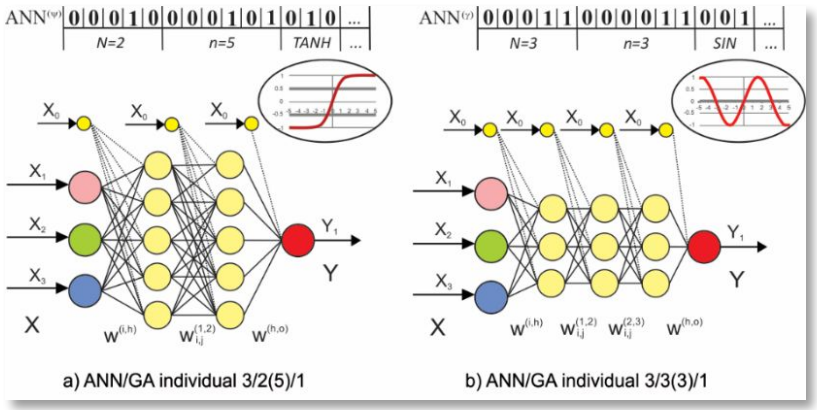
1. User Engagement & Ecosystem Reach (Annual Targets)
   - **Private AI Users:** Serving at least 75 individuals and 15 organizations annually, with a specific focus on 15 Startups and 15 SMEs.
   - **Public AI Users:** Serving 75 individuals and 15 public organizations annually.
   - **Ecosystem Breadth:** Supporting at least 200 participants within the broader AI Factory ecosystem, including European partners.

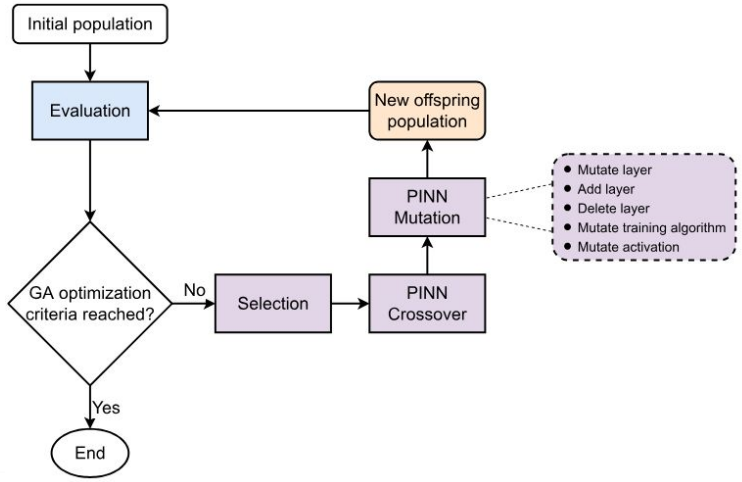
2. Human Capital & Training (Developing National Talent). Building a sustainable AI workforce:
   - **Individuals Trained:** A total of 500 individuals will receive specialized training.
   - **Student Participation:** Engaging 300 students in formal and informal AI Factory activities.
   - **Organization Training:** Providing training to 20 Startups and 20 SMEs to boost industrial AI adoption.
   - **Expert Capacity:** Maintaining 18 dedicated AI researchers and offering 10 specialized training programs.

FSUKG Engagement in SAIFA

Automated creation of ML models

Evolutionary algorithms.
Distributed computing.
Optimized hyperparameters.





Accelerate modeling process

Automated feature selection and scaling

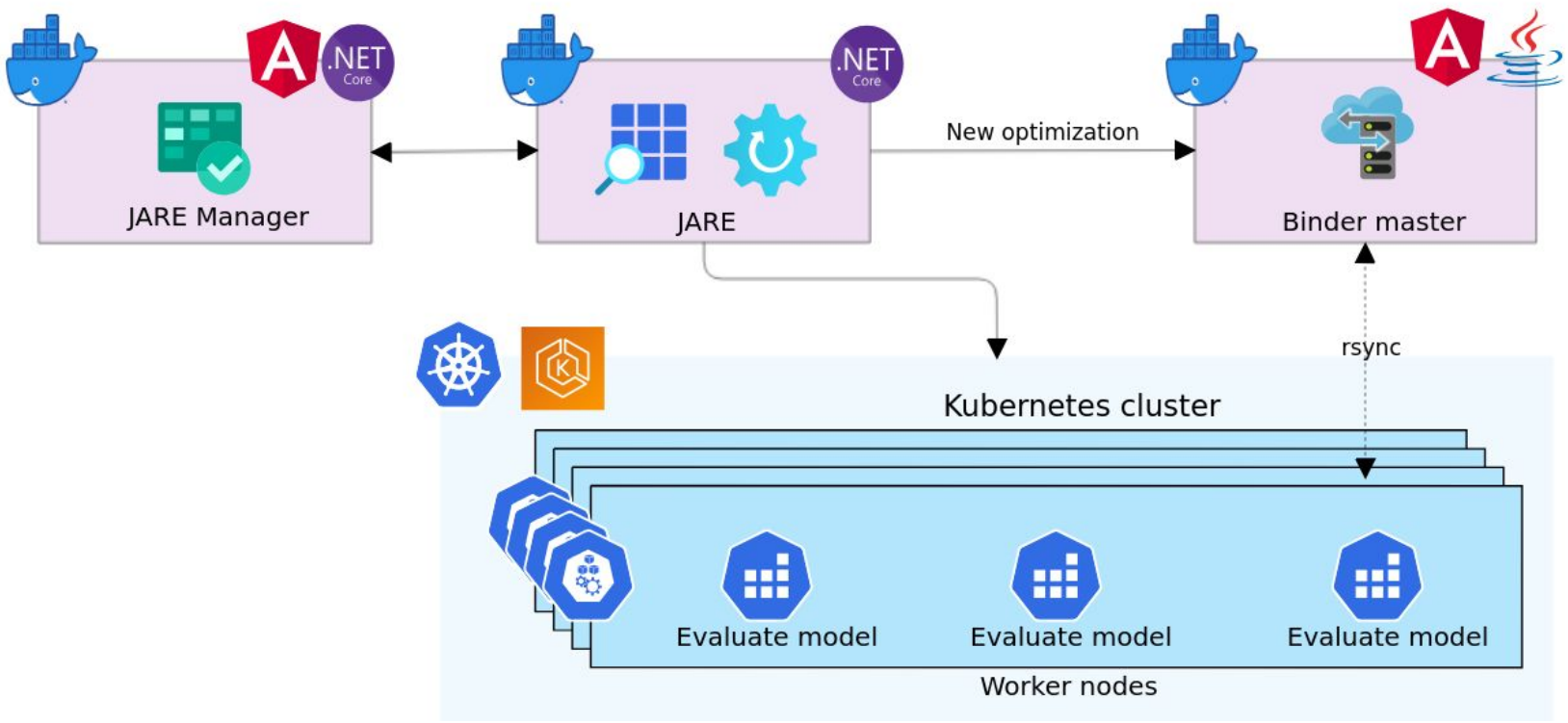
Optimized model architecture and parameters

BLACKFOX

SERBIAN ARTIFICIAL INTELLIGENCE FACTORY ANTENNA

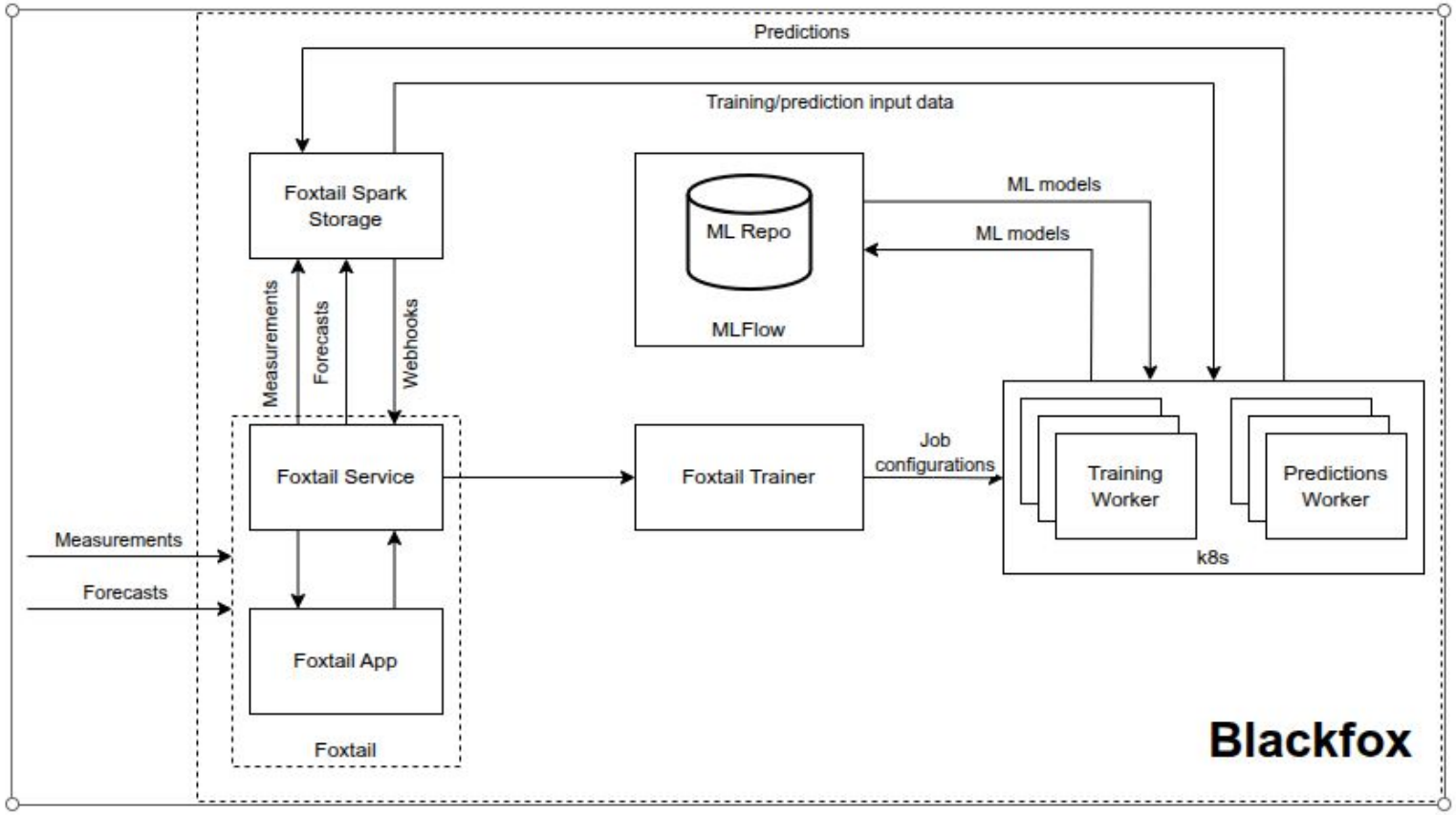
BlackFox Internals

- Based on microservices
- Available on HPC and Kubernetes
- Utilizes all available resources



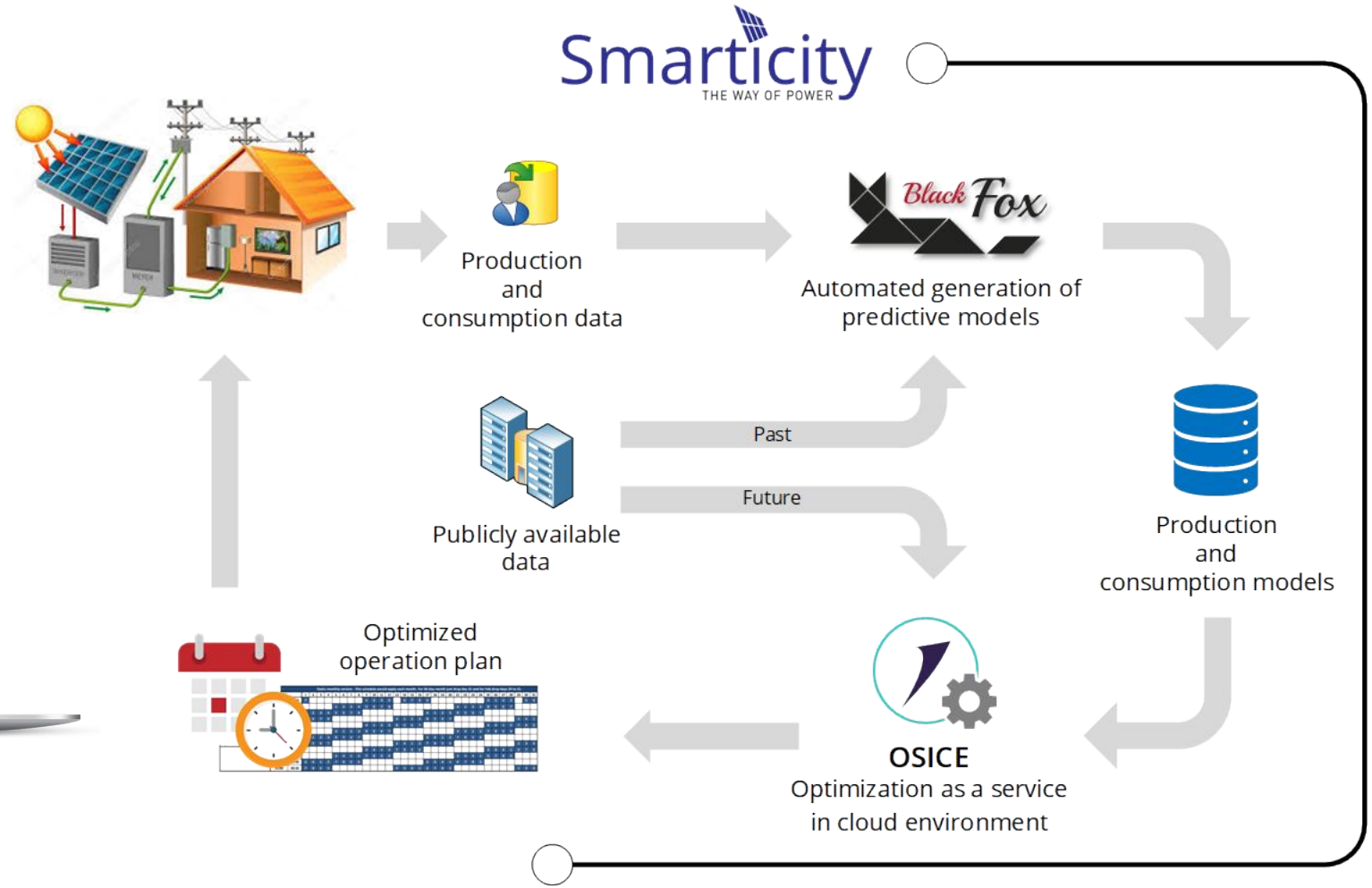
MLOps - Maintain 100s of AI Models

- ML Model repository based on MLFlow
- Asynchonous training triggered by events such as “new data available”
- Any data format available to Apache Spark



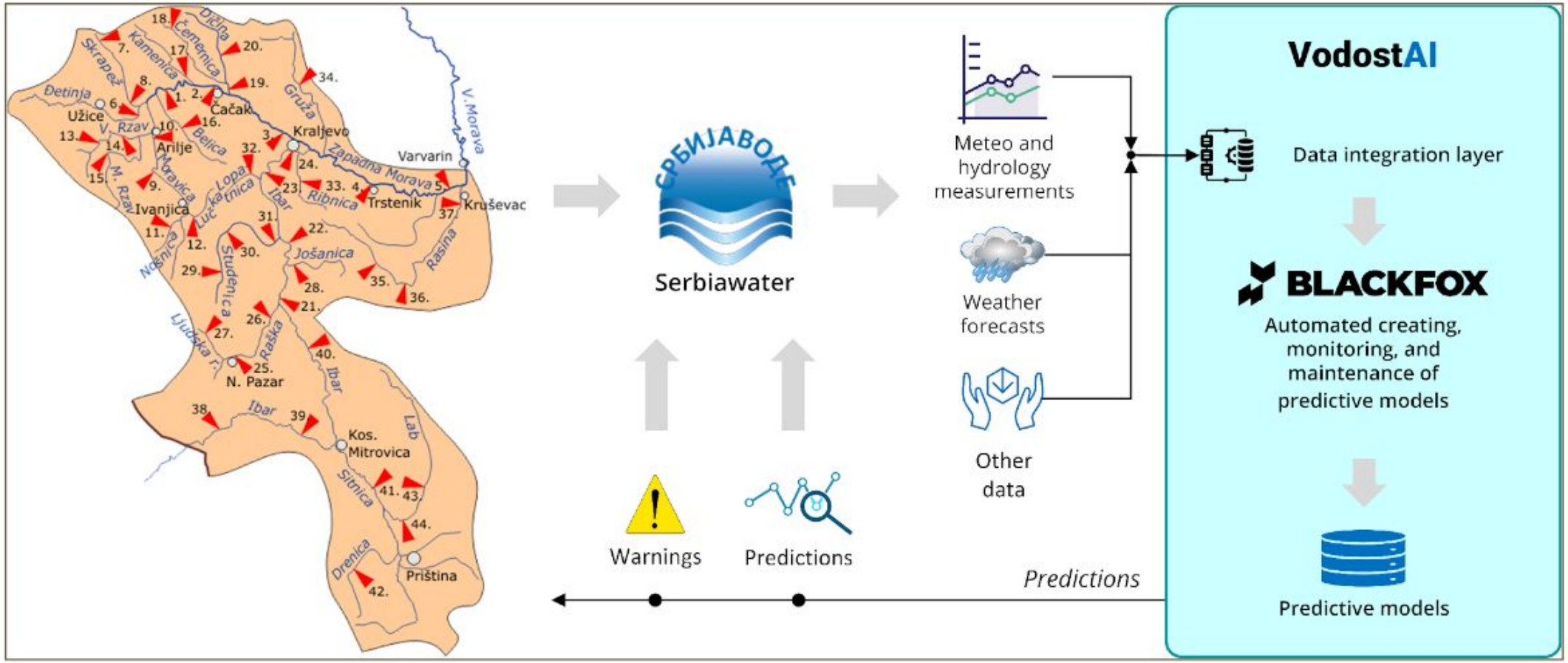
Smarticity

Smarticity automates finding an **optimal pattern in energy consumption and production** in case of facilities with renewable energy resources.



VodostAI

VodostAI is a software platform for **flood prediction and prevention** based on AI methods and IoT.



SAIFA - The First Project Year

- **Immediate Access to Supercomputing (Month 3):** Opening of user accounts on the PARADOX and ITE clusters, providing the computational power needed for high-level AI research and prototyping.
- **AI-Ready Datasets (Month 8):** Release of the first curated, high-quality datasets in priority sectors like Healthcare, Energy, and Environment, providing a foundation for local model training.
- **The SAIFA AI Gateway Beta (Month 12):** Launch of the national "one-stop-shop" web portal, simplifying how users request compute time, find tools, and access expert support.
- **First Open AI Models (Month 12):** Publication of initial open-source AI models across all priority sectors, demonstrating regional capabilities and providing templates for further innovation.
- **Launch of Specialized Training (Month 12):** Kick-off for the first formal training programs, including AI literacy courses and technical deep-dives designed to equip the workforce with exascale-ready skills.
- **Community Outreach Event (Month 12):** The first major national event to connect researchers, industry leaders, and policymakers, fostering a vibrant Serbian AI ecosystem.

SAIFA Is Just the Next Step

While SAIFA is a new initiative, it is effectively the "next chapter" of the roadmap Antun Balaž helped draft.

- **Validated Infrastructure:** The PARADOX cluster he built serves as the national validation hub where SAIFA models are developed and tested before transitioning to European AI Factories.
- **Human Capital:** His dedication to training and research excellence created the pool of dedicated AI researchers and technical experts who now form the core of the SAIFA consortium.
- **Sovereign AI Foundation:** His lifelong work ensured that Serbia possessed the local infrastructure and expertise necessary to participate as a lead partner in the EuroHPC JU, rather than being a mere consumer of foreign technology.

Thank You for Your Attention !