

# EuroHPC JU Information Day for AI on Supercomputers Dr. Lilit Axner Programme Officer at EuroHPC JU



# **Operational systems | Pre-exascale**

LUMI Consortium (Coordinator CSC) Kayaani, Finland

Leonardo Consortium (Coordinator CINECA) Bologna, Italy

Cray EX, Hewlett Packard Enterprise #3 Top500 (Nov 2023): **309.1** PFlops (LUMI-G)

AMD platform

- CPU: 64-core next-generation AMD EPYC<sup>™</sup>
- GPU: AMD Instinct<sup>™</sup> (MI250X),

Atos BullSequana XH2000 #4 Top500 (Nov 2022): 238.7 PFlops (BOOSTER)

> Intel/NVidia platform • CPU: Intel Sapphire Rapids • GPU: Nvidia custom Ambere (A100)

# **Operational systems | Petascale**

#### Vega



MeluXina



Karolina



Discoverer



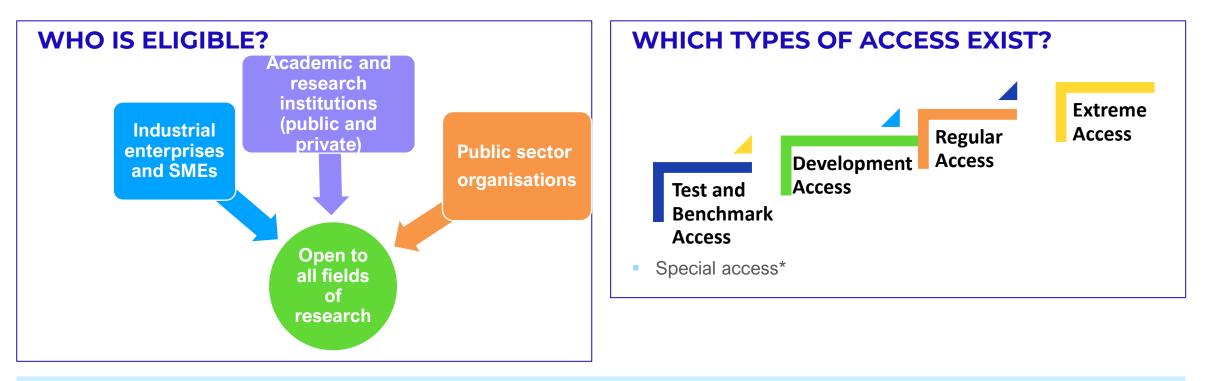
Sustained performance:	6,9 petaflops	Susta
CPU:	AMD Epyc Rome	CPU:
GPU:	Nvidia A100	GPU:
TOP500 ranking:	#32 in EU; #106 globally ( <u>June 2021</u> )	TOP5
Vendor/model	Atos BullSequana XH2000	Vend
Operated by	IZUM, Maribor, Slovenia	Opera

ustained performance:	12,8 petaflops
PU:	AMD Epyc Rome
PU:	Nvidia A100
OP500 ranking:	#10 in EU; #36 globally ( <u>June 2021</u> )
endor/model	Atos BullSequana XH2000
perated by	LuxProvide, Bissen, Luxembourg

Sustained performance:	9,13 petaflops	Sustained performance:	4,45 petaflops	
CPU:	AMD Epyc Rome	AMD Epyc Rome CPU:		
GPU:	Nvidia A100	GPU:	-	
TOP500 ranking:	#20 in EU; #69 globally ( <u>June 2021</u> )	TOP500 ranking:	#27 in EU; #91 globally ( <u>June 2021</u> )	
Vendor/model	HPE Apollo 2000Gen10 Plus and Apollo 6500	Vendor/model	Atos BullSequana XH2000	
Operated by	IT4I, Ostrava, Czech Republic	Operated by	PSB consortium, Sofia, Bulgaria	

## **Access to EuroHPC Supercomputers**





### WHAT ARE THE CONDITIONS FOR ACCESS?

Access is free of charge. Participation conditions depend on the specific access call that a research group has applied to. In general users of EuroHPC systems commit to:

- acknowledge the use of the resources in their related publications
- contribute to dissemination events
- produce and submit a report after completion of a resource allocation

## **GPU Node Hours Provided per Call type**



Access type	GPU Node hours*	Duration
Benchmark access	Between 400 to 3000	3 months
Development access	Between 1000 and 10 000	6 or 12 months
Regular access	Up to 730 000	12 months
Extreme access	Up to 7.7 million	12 or 24 months

### Node hours = Number of Nodes x Number of Cores per Node

#### **Documentation on Current systems with GPU partitions**

https://doc.vega.izum.si/ - VEGA https://docs.it4i.cz/karolina/hardware-overview/ -Karolina https://docs.lxp.lu/ - Meluxina https://docs.lumi-supercomputer.eu/ - Lumi https://leonardo-supercomputer.cineca.eu/hpc-system/ - Leonardo



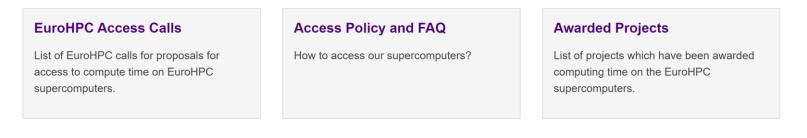


### The European High Performance Computing Joint Undertaking (EuroHPC JU)



Home > Access to Our Supercomputers

#### Access to Our Supercomputers



#### Share this page



https://eurohpc-ju.europa.eu/access-our-supercomputers/eurohpcaccess-calls\_en



Search

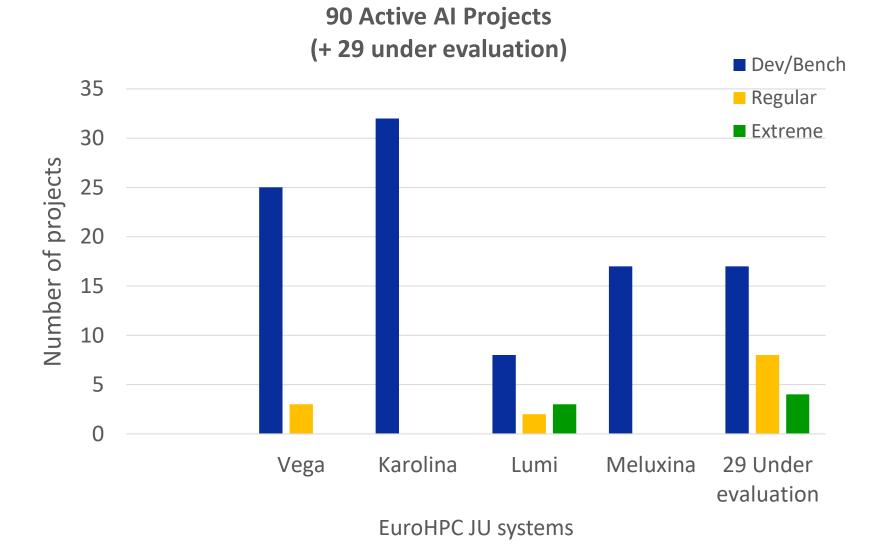


Film produced by ENCCS



### AI Applications per System by August 2023

(42 of 119 are LLM applications)



## **Expert Support**

**EuroHPC JU Hosting Entities** 





#### 33 EuroCC National Competence Centres across Europe



# Support @ EuroHPC JU Hosting Entities

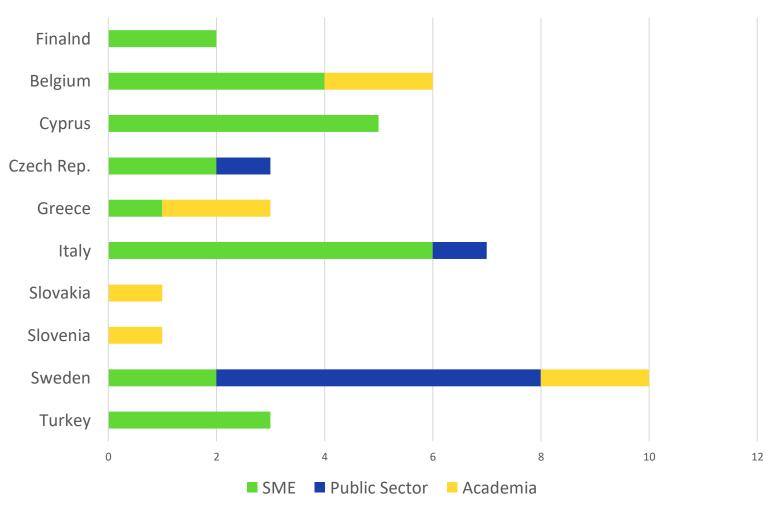


Centre	Projects	FTEs	Software & Tools	Training
CINECA	91	2*	PyTorch, Tensorflow, Keras, NetKet, Jax, Keras, Caffe, OpenCV, Weights&Biases,	<ul> <li>✓ Introduction to Machine and Deep Learning</li> <li>✓ Fundamentals of Deep Learning</li> <li>✓ Fundamentals of Deep Learning for Multi-GPUs</li> </ul>
IT4I	99	10	Tensorboard, Torchmetrics, Torchtext, Torchvision, Transformers, LLama, OpenGPT-X, Xgboost, Horovod, Deepspeed,	<ul> <li>✓ Fundamentals of Deep Learning for Multiple Data Types</li> <li>✓ Fundamentals of Deep Learning for Computer Vision</li> </ul>
CSC	14	3	Dm-tree, Lightning-bolts, Onnx, Pyarrow, Scikit-learn, Scipy, Spacy, Ray, Megathron	<ul> <li>✓ Practical Deep Learning course</li> <li>✓ Data Science and Deep Learning with Python</li> <li>✓ Data science with R</li> </ul>
				<ul> <li>✓ Introduction to Deep Learning and TensorFlow</li> <li>✓ Seasonal School of Computer Graphics for Cultural Heritage: AI and Cultural Heritage</li> </ul>



## Support @ your country by EuroCC NCCs

<u>https://www.eurocc-</u>
 <u>access.eu/about-us/meet-</u>
 <u>the-nccs/</u>



#### AI projects supported by EuroCC NCCs in their countries

# **EuroCC Training on Al**



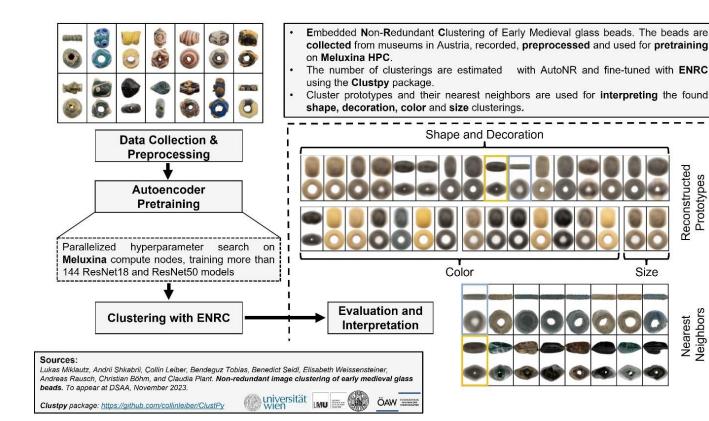
- ✓ High Performance Data Analytics in Python
- ✓ Data Analysis and Plotting in Python with Pandas
- ✓ Advanced Deep Learning with Transformers
- ✓ Megatron Bootcamp
- Training on HPDA for climate data with the Ophidia framework
- ✓ Upscaling AI with Containers
- ✓ Applications of AI for Predictive Maintenance
- ✓ Applications of AI for Anomaly Detection
- ✓ Fundamentals of Deep Learning for Multi-GPUs
- ✓ Efficient multi-GPU and multi-node execution of AI applications and frameworks
- ✓ AI for Industry
- ✓ AI on Supercomputers
- ✓ AI for Industry: Advance level

- ✓ AI basics
- ✓ Practical Deep Learning workshop
- ✓ "AI for Science" NVIDIA Bootcamp
- ✓ Workshop on High-performance Data Analytics
- ✓ Al as a Tool for Change
- ✓ Advanced Deep Learning with Transformers
- ✓ Introduction to Deep-Learning
- ✓ Creative AI Webinar

https://www.eurocc-access.eu/services/training/

## Al to Cluster Medieval Glass Beds -MeluXina



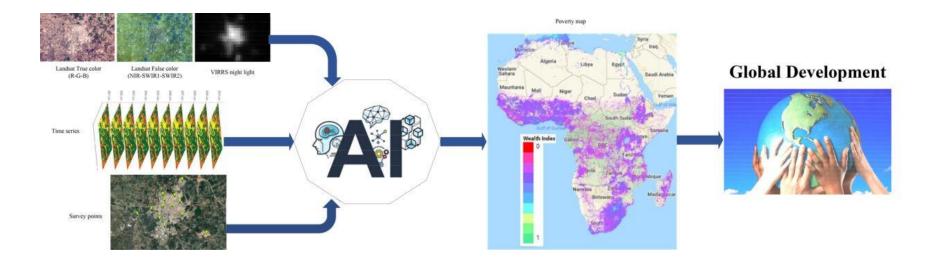


Glass beads are among the most common grave goods in the early Middle Ages, and their number can be estimated in the millions. The colour, size, shape, production technique and decoration of the beads that are discovered in burial sites contain much information that is relevant for historians regarding social customs, trade routes and production networks.

A research group from the University of Vienna sought to improve and validate the accuracy of existing deep embedded non-redundant clustering methods to find different informative ways to categorise the glass beads.

## Al for Poverty Mitigation - Karolina





A research group from the University of Gothenburg wanted to better understand the distribution of global poverty historically and geographically. To do this, they set out to train deep-learning models to predict health and living conditions using satellite images.

The research group used TensorFlow on the EuroHPC JU Karolina supercomputer in the Czech Republic to address their project and successfully implement their observations.

The new data gathered will allow scholars to examine the causal effects of foreign aid on the likelihood of impoverished communities overcoming poverty. This, in turn, will enhance the alignment of development and aid initiatives with the challenges they aim to address.

# Using EuroHPC Vega System by the Swedish National Archives

### Vega for training and inference

- Training the SATRN-model on Vega enabled us to increase the scale of the resized images going into the model, thereby improving accuracy for handwritten text, which generally requires more information than printed text
- Running 9 million images thorugh the pipeline on VEGA took roughly 90 node-hours
- At a hit-rate of 90% this project saves us about 700000 euros in manual labor costs, and the indexing database gets created a lot quicker



Adapting Altechnology for use in archives

- Image segmentation models
- Text-recognition

Make scanned images searchable

384 000 GPU core hours (Development Access, VEGA)

### The Property Record Indexing Pipeline

