interface

A Critical Appraisal of the EU AI Compute Strategy – Where is the Demand?

March 19, 2026 **Workshop The EU strategy on AI: Ensuring autonomy and prosperity @ *arbeit.digital Zukunftskonferenz***

All eyes on compute

What we know for sure: AI Compute seen as the key to becoming a leading European 'AI continent'.

Thus, the European Commission plans a massive expansion of AI compute, via

- the **Cloud & AI Action Plan** (3x more datacentre capacity by 2030) → with the goal to boost technological sovereignty
- the creation of a network of 19 **AI Factories** across Europe, each equipped with up to 25.000 H100 GPU equivalents → with the goal to provide small and medium-sized enterprises (SMEs), researchers, and startups with access to the computing power needed to develop and test AI systems
- and **5 AI Gigafactories** (large GPU clusters) with 100.000 H100 GPUs equivalents → with the goal to …?

Julia Hess - jhess@interface-eu.org

Zooming in on AI Gigafactories (AIGFs): Where we are in the process

<u>Key Commission / EuroHPC actions:</u>

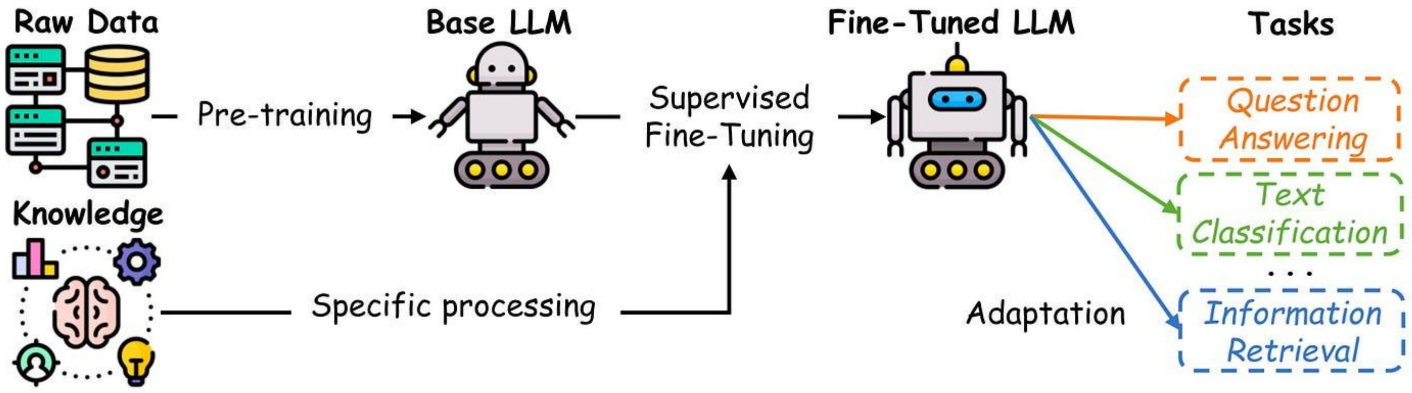
- Issue a call for expression of interest to invest in AI Gigafactories (9 April 2025);
- Define the InvestAI Facility with EIBG (Q3/Q4 2025);
- Launch the official call on AI Gigafactories under the EuroHPC Joint Undertaking (Q4 2025);
- Address the financing gap of startups and scaleups and facilitate their access to markets,–public procurement, services and talent in the EU Startup and Scaleup Strategy (Q2 2025).

Still waiting for the official call, criteria for the AIGFs

Julia Hess - jhess@interface-eu.org

Zooming in on AIGFs: Phases of Generative AI Training and Deployment

1. **Pre-training:** models learn general patterns from very large datasets using self-supervised learning.
2. **Post-training:** the pre-trained model then is refined to improve performance for specific goals, either via **Fine-tuning** (training the model on domain-specific data (e.g., medical literature)) or **Reinforcement learning** (adjusting the model's behaviour based on human or automated feedback)
3. **Inference:** Finally, the model is deployed in user-facing applications, such as chatbots or search engines. At this stage, additional computational resources – known as *inference compute* – are required each time the model generates an output.

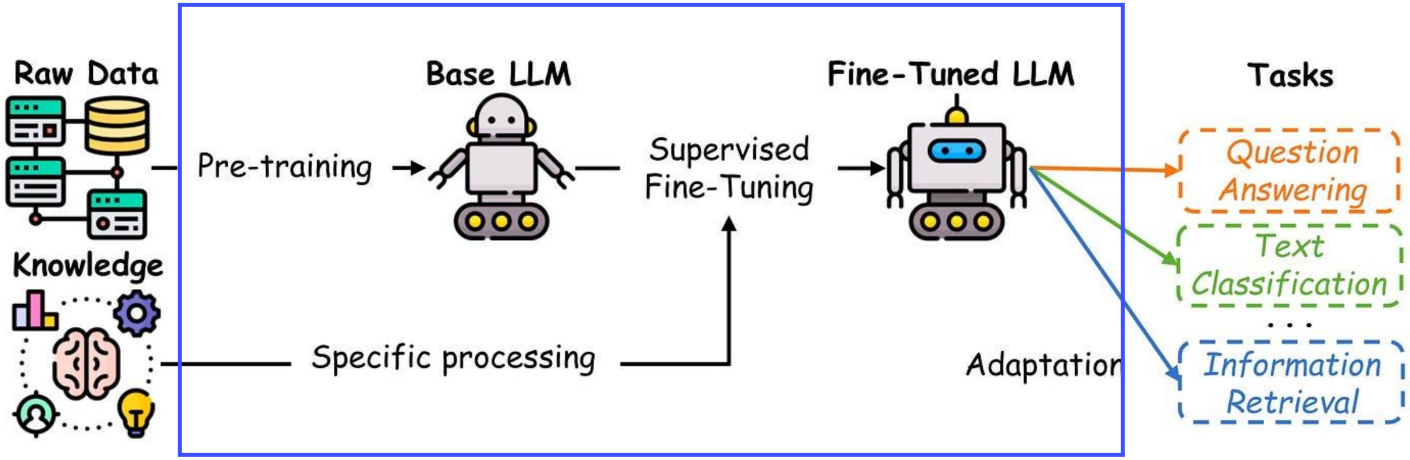


Julia Hess - jhess@interface-eu.org

interface

Purpose of the AIGFs?

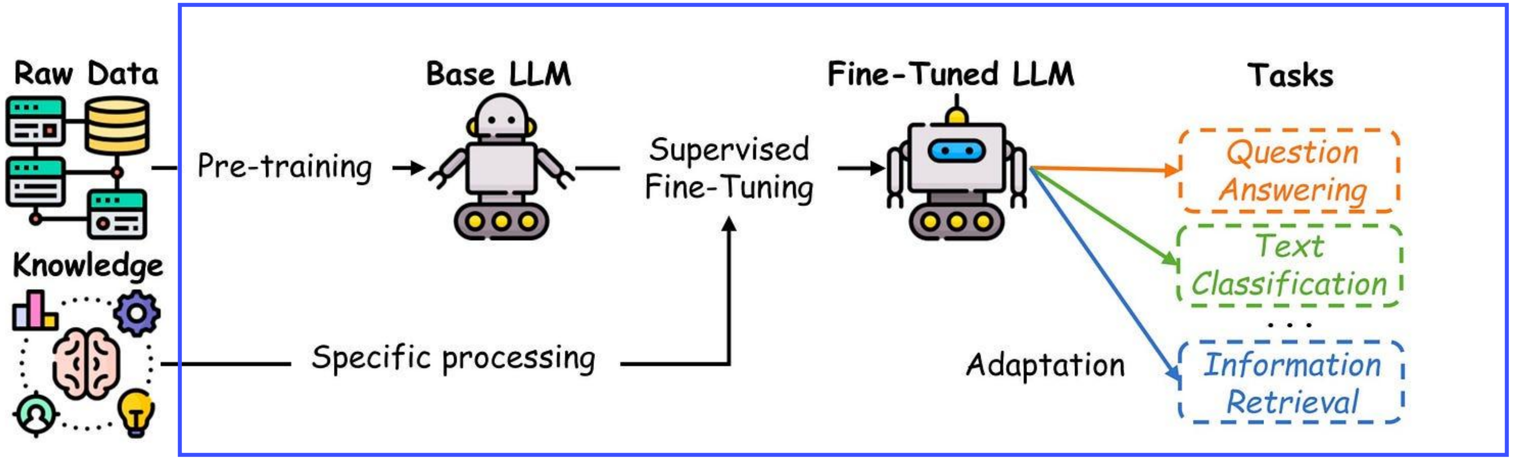
AI Continent Action Plan (April 2025): *"AI Gigafactories will be large-scale facilities that develop and train complex AI models at an unprecedented scale, with hundreds of trillions of parameters. They will integrate massive computing power, exceeding 100,000 advanced AI processors, while taking into account power capacity, as well as energy, water efficiency and circularity. These facilities are essential for Europe to compete globally and maintain its strategic autonomy in scientific progress and critical industrial sectors. They will be federated with the EuroHPC network of AI Factories, ensuring seamless integration and knowledge sharing across the European AI ecosystem."*



Julia Hess - jhess@interface-eu.org

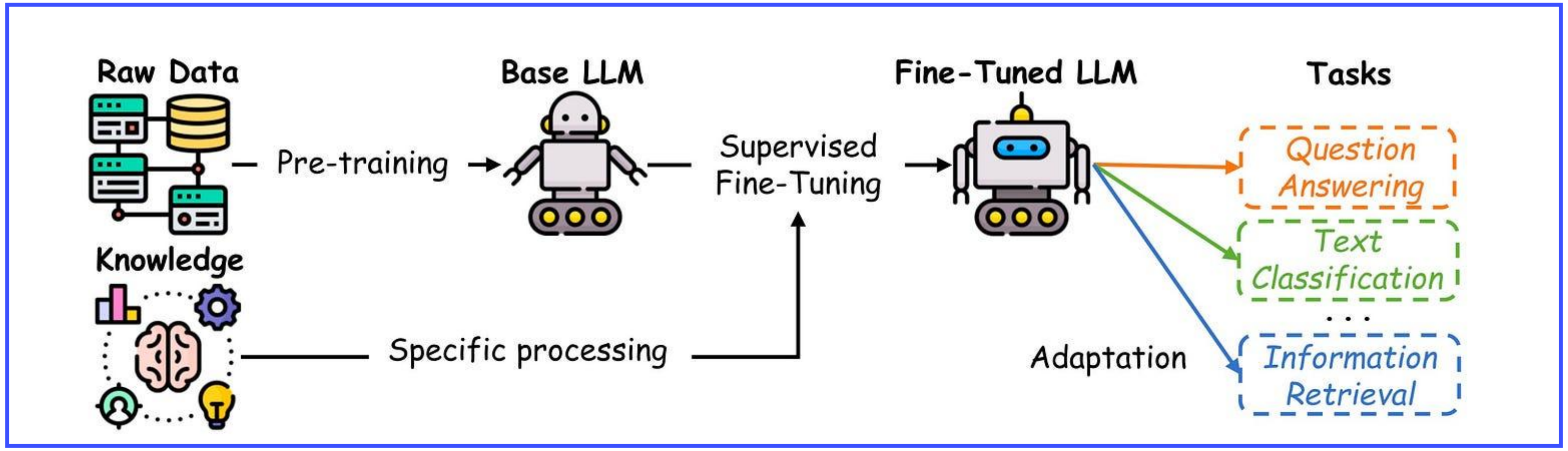
Purpose of the AIGFs?

Call for Expression of Interest in AI Gigafactories (April 2025): *"AI Gigafactories (AIGFs) will build on the concept of AI Factories and take it to the next level. These will be large-scale AI compute infrastructure facilities (including the required data storage infrastructure facilities) designed to develop, train, and deploy very large AI models and applications at an unprecedented scale (e.g., AI models in the order of hundreds of trillions of parameters). They will be integrating massive computing power (e.g., beyond 100,000 advanced AI chips compared to 25,000 in the largest AI Factory's supercomputers), together with energy-efficient data centres, and AI-driven automation to optimise AI model training, inference, and deployment."*



Purpose of the AIGFs?

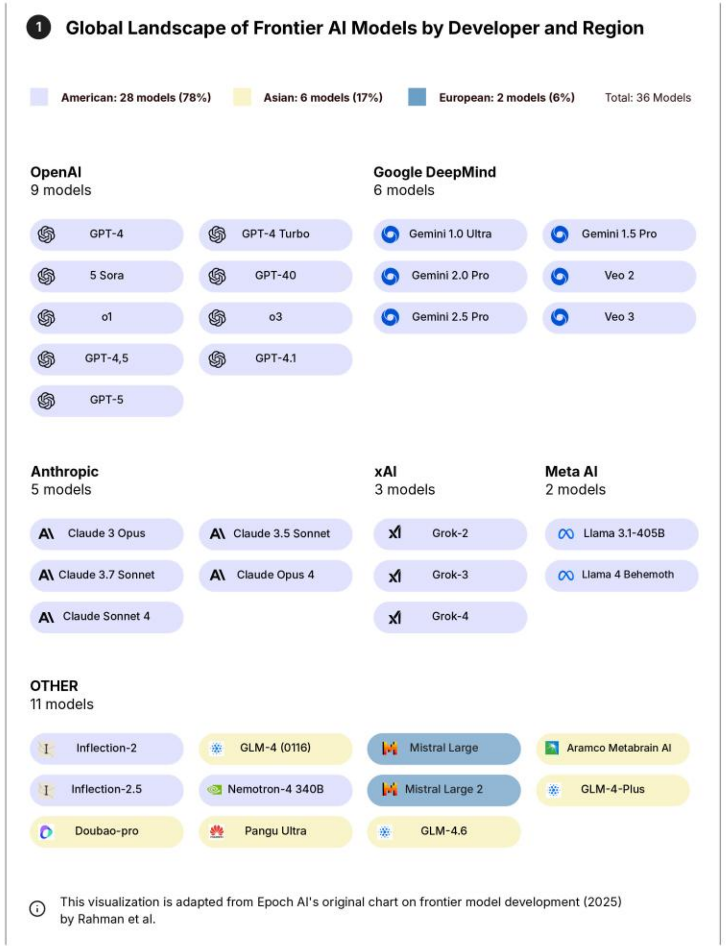
Updated HPC Regulation (January 2026): *"Artificial Intelligence gigafactory" or "AI gigafactory" means a state-of-the-art large-scale facility with sufficient capacity to handle the complete lifecycle, from development to large-scale inference, of very large AI models and applications, providing a supercomputing service infrastructure which is composed of AI-optimised computing capacity, a supporting data centre infrastructure including high-capacity storage and networking, dedicated secure cloud user access environments, and specialised secure AI-oriented support services for its advanced operations, all of which are supported by an environmentally sustainable infrastructure, in particular for energy and water supply systems."*



= no agreement on the purpose!

interface

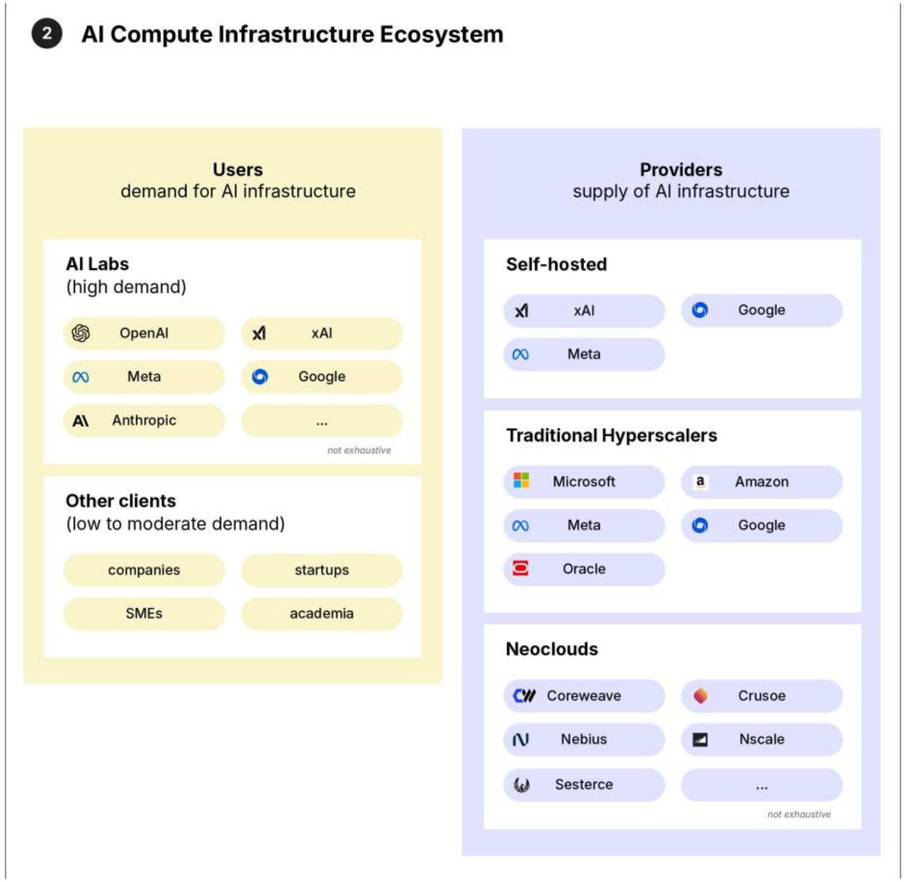
1. Mapping AI Labs: Who is developing today's frontier models?



- 36 frontier AI models worldwide
  - 78% US
  - 17% Asia
  - 6% Europe

- Only a small number of AI labs are developing today's frontier models
  - 25 of 36 models come from just five AI labs, with OpenAI alone responsible for nine (OpenAI, Google DeepMind, Anthropic, xAI, Meta AI)
  - EU: only Mistral (FR)

2. Mapping the User-Provider-Landscape & Capacity Expansion: Who is driving the demand?

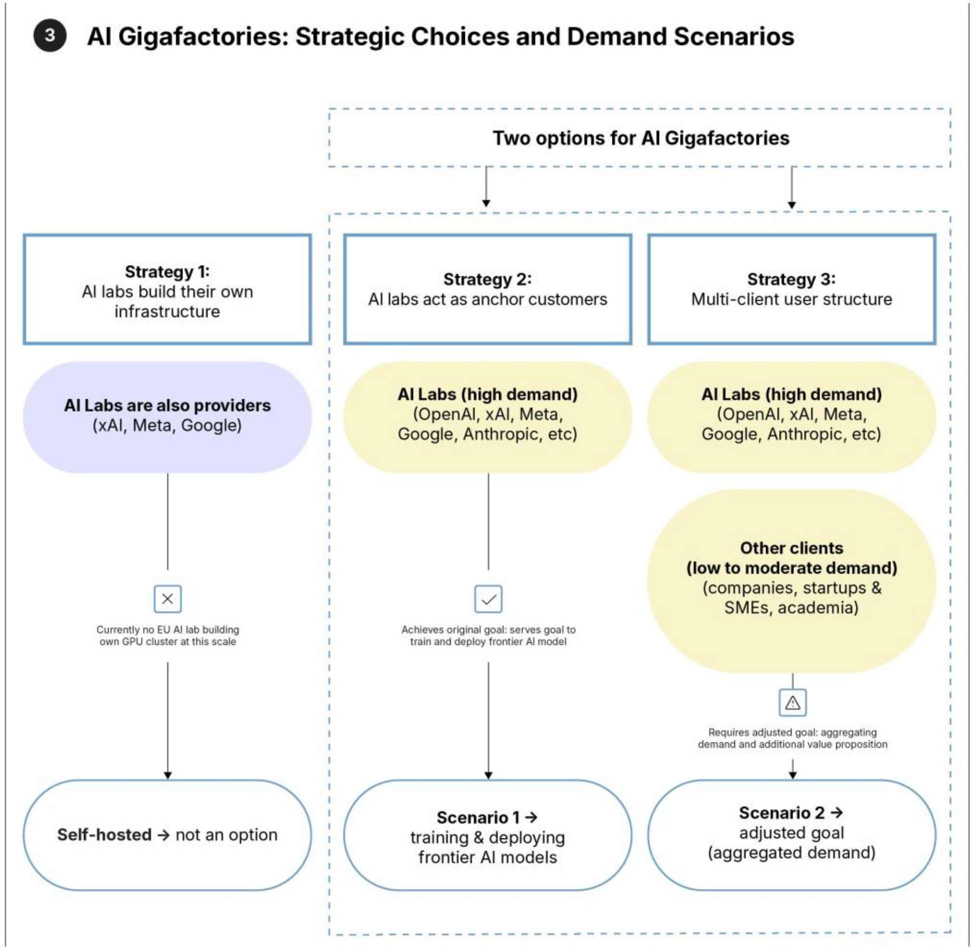


Country	Location	Name	Available Since/ From	Provider	User	H100 equivalents
US	Omaha, Nebraska	**Papillion Campus**	2021	Google	Google	?
US	Omaha, Nebraska	**Google Omaha AI Datacenter**	2024	Google	Google	?
US	Mountain View, California	**Lambda Labs Cluster**	2023	Lambda	multi-tenant	?
US	Lawrence, Kansas	**Lawrence Livermore NL El Capitan Phase 2**	2024	US Department of Energy	multi-tenant	?
US	Austin, Texas	**Tesla Cortex Phase 1**	2024	Tesla	Tesla	50 000
US	Undisclosed Location	**Meta AI Research Supercluster (RSC)**	2024	Meta	Meta	2× 24 000
France	Paris	**Nebius Paris cluster**	2024	Nebius	multi-tenant	?
US	Memphis, Tennesee	**Colossus Phase 1 (xAI)**	2024	xAI	xAI	100 000

You can find all entries in our publication!

interface

3. Takeaways for AIGFs: **Anchor Customers vs. Multi-Client User Structure**



- **Frontier AI labs are only user group with large workloads** to “train and deploy frontier AI”, they either operate their own infrastructure (e.g., xAI) or act as anchor customers (e.g., OpenAI) → see Strategy 1 & 2

  - **Anchor Customer Strategy in EU only possible with Mistral** (see their announcements with Fluidstack)

  - **Alternative Multi-Client User Structure**: aggregate sufficient demand from clients with low-to-moderate AI workloads similiar to neoclouds (purpose **no longer to train and deploy frontier AI models**, need new value proposition**) → See strategy 3**

Julia Hess - jhess@interface-eu.org

Checklist for AIGFs

Demand Quantification

Demand estimates should be a core requirement of each submission. Given the scarcity of reliable public data, proposals should include robust projections of AI workloads to be served, supported by concrete evidence of demand (e.g. letters of intent or pre-commitments).

Key questions:

1. Who is driving AI compute demand in Europe?
2. Which stakeholders require significant compute capacity (e.g. industry, research, public sector)?
3. How much demand originates from ministries and critical sectors, where sovereignty considerations may be more important?
4. What is the expected distribution between training and inference workloads?

Checklist for AIGFs

A Realistic Strategic Objective

Strengthening the European AI ecosystem requires more than scaling compute capacity. It requires a clear strategic objective that is aligned with industry dynamics and Europe's comparative advantages.

Key questions:

1. How are existing policy initiatives connected (e.g. Cloud and AI Development Act, Apply AI Strategy, Member State compute strategies, Frontier AI initiatives, AIFs, AIGFs)?
2. If the full AI lifecycle is being addressed, what distinct roles do AI Factories (AIFs) and AI Gigafactories (AIGFs) play, and how do they complement each other?
3. What is the long-term strategy that reflects the structure and strengths of the European AI ecosystem?
4. How can this strategy remain adaptable to rapid technological change and shifting geopolitical conditions?

Checklist for AIGFs

A Clear Value Proposition

The success of AI Gigafactories (AIGFs) depends on a clear positioning and differentiation from existing providers, particularly hyperscalers and emerging neoclouds.

Key questions:

1. Beyond providing sovereign compute, how can AIGFs create unique value for the European AI ecosystem?
2. In which use cases or sectors do AIGFs offer a comparative advantage?
3. What are the benefits of choosing an AIGF over an EU-based neocloud provider?